

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 14, 2000, 10:06:42 ; Search time 72.12 Seconds  
(without alignments)  
642.433 Million cell updates/sec

Title: US-08-978-456-2  
Perfect score: 1789  
Sequence: 1 MDVAIQLANMVGGGTGVNPP.....FEIVHSELLNQNVLTKRK 342

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 866916 seqs, 135474527 residues

Total number of hits satisfying chosen parameters: 866916

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending Patents-AA:\*  
1: /cgn2\_6/ptodata/2/paa/PCTUS\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/paa/US06\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/paa/US07\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/paa/US080\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/paa/US081\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/paa/US082\_COMB.pep.\*  
7: /cgn2\_6/ptodata/2/paa/US083\_COMB.pep.\*  
8: /cgn2\_6/ptodata/2/paa/US084A\_COMB.pep.\*  
9: /cgn2\_6/ptodata/2/paa/US084B\_COMB.pep.\*  
10: /cgn2\_6/ptodata/2/paa/US085\_COMB.pep.\*  
11: /cgn2\_6/ptodata/2/paa/US086\_COMB.pep.\*  
12: /cgn2\_6/ptodata/2/paa/US087\_COMB.pep.\*  
13: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep.\*  
14: /cgn2\_6/ptodata/2/paa/US089\_COMB.pep.\*  
15: /cgn2\_6/ptodata/2/paa/US090\_COMB.pep.\*  
16: /cgn2\_6/ptodata/2/paa/US091\_COMB.pep.\*  
17: /cgn2\_6/ptodata/2/paa/US092\_COMB.pep.\*  
18: /cgn2\_6/ptodata/2/paa/US093\_COMB.pep.\*  
19: /cgn2\_6/ptodata/2/paa/US094\_COMB.pep.\*  
20: /cgn2\_6/ptodata/2/paa/US095\_COMB.pep.\*  
21: /cgn2\_6/ptodata/2/paa/US096\_COMB.pep.\*  
22: /cgn2\_6/ptodata/2/paa/PCT\_NEW\_COMB.pep.\*  
23: /cgn2\_6/ptodata/2/paa/US05\_NEW\_COMB.pep.\*  
24: /cgn2\_6/ptodata/2/paa/US07\_NEW\_COMB.pep.\*  
25: /cgn2\_6/ptodata/2/paa/US08\_NEW\_COMB.pep.\*  
26: /cgn2\_6/ptodata/2/paa/US09\_NEW\_COMB.pep.\*  
27: /cgn2\_6/ptodata/2/paa/US60\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1789	100.0	342	1	PCT-US98-24857-2
2	1759.5	98.4	372	13	US-08-827-356-4163
3	1759.5	98.4	372	20	US-09-611-529-5797
4	1193.5	66.7	350	18	US-09-450-969-6115
5	500	27.9	379	27	US-60-215-161-5966
6	485.5	27.1	366	18	US-09-328-352-5462

7	469	26.2	366	19	US-09-583-110-4805	Sequence 4805, Ap
8	467	26.1	366	1	PCT-US98-25010-2	Sequence 2, Appli
9	467	26.1	366	14	US-08-979-616-2	Sequence 2, Appli
10	464	25.9	402	17	US-09-252-691-7588	Sequence 7588, Ap
11	451	25.2	99	1	PCT-US97-02318-354	Sequence 354, App
12	451	25.2	99	1	PCT-US98-24857-4	Sequence 4, Appli
13	451	25.2	99	14	US-08-903-470-354	Sequence 354, App
14	442.5	24.7	387	16	US-09-107-532-5675	Sequence 5675, Ap
15	435	24.3	378	26	US-09-438-185-873	Sequence 873, App
16	404	22.6	352	19	US-09-540-236-3272	Sequence 3272, Ap
17	404	22.6	352	21	US-60-128-476-4795	Sequence 4795, Ap
18	393	22.0	410	15	US-09-064-935-15	Sequence 15, Appli
19	356	19.9	401	18	US-09-369-953-5	Sequence 5, Appli
20	334.5	18.7	407	20	US-09-614-912-78	Sequence 78, Appli
21	334.5	18.7	407	21	US-60-170-906-10	Sequence 10, Appli
22	320	17.9	599	21	US-60-170-906-15	Sequence 15, Appli
23	300	16.8	363	21	US-60-170-906-14	Sequence 14, Appli
24	242	13.5	288	20	US-09-614-912-80	Sequence 80, Appli
25	242	13.5	288	21	US-60-170-906-12	Sequence 12, Appli
26	170.5	9.5	138	27	US-60-215-161-5704	Sequence 5704, Ap
27	153	8.6	155	1	PCT-US97-14436-650	Sequence 650, App
28	153	8.6	155	14	US-08-911-503-650	Sequence 650, App
29	153	8.6	155	14	US-08-911-503A-650	Sequence 650, App
30	151	8.4	140	19	US-09-583-110-4436	Sequence 4436, Ap
31	148.5	8.3	169	18	US-09-328-352-7583	Sequence 7583, Ap
32	143.5	8.0	173	16	US-09-134-000-4901	Sequence 4901, Ap
33	140	7.8	333	14	US-08-993-002A-8132	Sequence 8132, Ap
34	140	7.8	344	1	PCT-US98-06371-1112	Sequence 1112, Ap
35	140	7.8	344	14	US-08-902-615A-378	Sequence 378, App
36	140	7.8	344	14	US-08-993-002A-8133	Sequence 8133, Ap
37	136.5	7.6	171	16	US-09-107-532-4012	Sequence 4012, Ap
38	136	7.6	345	14	US-08-993-002A-8134	Sequence 8134, Ap
39	133.5	7.5	173	18	US-09-450-969-5242	Sequence 5242, Ap
40	128.5	7.2	158	3	US-07-928-506-8	Sequence 8, Appli
41	128.5	7.2	158	3	US-07-928-506A-9	Sequence 9, Appli
42	128.5	7.2	158	4	US-08-066-297-23	Sequence 23, Appli
43	127.5	7.1	158	1	PCT-US99-10985-16	Sequence 16, Appli
44	125.5	7.0	158	1	PCT-US99-10985-14	Sequence 14, Appli
45	123	6.9	187	17	US-09-252-691-6455	Sequence 6455, Ap

ALIGNMENTS

RESULT 1  
PCT-US98-24857-2  
; Sequence 2, Application PC/TUS9824857A  
; GENERAL INFORMATION:  
; APPLICANT: SmithKline Beecham Corporation  
; TITLE OF INVENTION: Novel ribG  
; FILE REFERENCE: P50444-9  
; CURRENT APPLICATION NUMBER: PCT/US98/24857A  
; CURRENT FILING DATE: 1998-11-23  
; EARLIER APPLICATION NUMBER: 08/978,456  
; EARLIER FILING DATE: 1997-11-25  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 2  
; LENGTH: 342  
; TYPE: PRT  
; ORGANISM: Staphylococcus aureus  
PCT-US98-24857-2

Query Match 100.0%; Score 1789; DB 1; Length 342;  
Best Local Similarity 100.0%; Pred. No. 1.8e-177;  
Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MDVAIQLANMVGGGTGVNPPGAVVNVNIGRVTGIGLHRLKGDHAEVQALDMAQNAEGA	60
Db	1	MDVAIQLANMVGGGTGVNPPGAVVNVNIGRVTGIGLHRLKGDHAEVQALDMAQNAEGA	60
Qy	61	TIYITLPCSHFGSTPPCNKIIDCKIAKVYATKNSLDTGHGDETLEHAGIEVCVDDE	120

Db 61 TTYITLPCSHFGSTPPCVNKKIIDCKIAKVYATKNSLDTHGDETLRAHGEIEVCVDD 120  
QY 121 RASQLYQDFFKAKAKOLPOITVKVSASLDGKQANDNGSQOWITNKVKQDVYKLRHRHDA 180  
Db 121 RASQLYQDFFKAKAKOLPOITVKVSASLDGKQANDNGSQOWITNKVKQDVYKLRHRHDA 180  
QY 181 VLTGRTVELDDPOYTTRIQQDGKNPIKVKLSKSGNIHFNOQIYQDESTPIWIYTENPNLT 240  
Db 181 VLTGRTVELDDPOYTTRIQQDGKNPIKVKLSKSGNIHFNOQIYQDESTPIWIYTENPNLT 240  
QY 241 SNOTHIEIILYKSCDLTTLILHNLKRGVGTLLVEAGPTTTSFSEIYYIDEFILYYAPKLI 300  
Db 241 SNOTHIEIILYKSCDLTTLILHNLKRGVGTLLVEAGPTTTSFSEIYYIDEFILYYAPKLI 300  
QY 301 GSGSNYQFYQTNVDVIEIPDANQFEIVHSELNQNVLTLRKK 342  
Db 301 GSGSNYQFYQTNVDVIEIPDANQFEIVHSELNQNVLTLRKK 342

## RESULT 2

US-08-827-356-4163  
; Sequence 4163 Application US/08827356  
; GENERAL INFORMATION:

; APPLICANT: George H. Shimer, Jr.  
; APPLICANT: George H. Miller  
; APPLICANT: Roberta S. Hare  
; APPLICANT: Karen J. Shaw  
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS RELATED  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS  
; NUMBER OF SEQUENCES: 5574  
; CORRESPONDENCE ADDRESSES:

; ADDRESSEE: Schering-Plough Corporation  
; STREET: 2000 Galloping Hill Road  
; CITY: Kenilworth  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07033-0530

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: PC-DOS/MS-DOS  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/827,356  
; FILING DATE: 01-APR-1997

; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/014,477  
; FILING DATE: 01-APR-1996  
; APPLICATION NUMBER: 60/016,743  
; FILING DATE: 02-MAY-1996  
; APPLICATION NUMBER: 60/020,016  
; FILING DATE: 14-JUN-1996

; INFORMATION FOR SEQ ID NO: 4163:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 372 amino acids  
; TYPE: amino acid

; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: YES  
; ORIGINAL SOURCE:  
; ORGANISM: Staphylococcus aureus  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: 1...372  
US-08-827-356-4163

Query Match 98.4%; Score 1759.5; DB 13; Length 372;  
Best Local Similarity 98.8%; Pred. No. 2,4e-174;  
Matches 339; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 MDYAIQLANWVGQTCVNPVPGAVVYVNEGRIVGIGHLRKGDHAEVQALDMAQQAEGA 60

Db 30 MDYAIQLANWVGQTCVNPVPGAVVYVNEGRIVGIGHLRKGDHAEVQALDMAQQAEGA 89  
QY 61 TTYITLPCSHFGSTPPCVNKKIIDCKIAKVYATKNSLDTHGDETLRAHGEIEVCVDD 120  
Db 90 TTYITLPCSHFGSTPPCVNKKIIDCKIAKVYATKNSLDTHGDETLRAHGEIEVCVDD 149  
QY 121 RASQLYQDFFKAKAKOLPOITVKVSASLDGKQANDNGSQOWITNKVKQDVYKLRHRHDA 180  
Db 150 RASQLYQDFFKAKAKOLPOITVKVSASLDGKQANDNGSQOWITNKVKQDVYKLRHRHDA 209  
QY 181 VLTGRTVELDDPOYTTRIQQDGKNPIKVKLSKSGNIHFNOQIYQDESTPIWIYTENPNLT 240  
Db 210 VLTGRTVELDDPOYTTRIQQDGKNPIKVKLSKSGNIHFNOQIYQDESTPIWIYTENPNLT 269  
QY 241 SNOTHIEIILYKSCDLTTLILHNLKRGVGTLLVEAGPTTTSFSEIYYIDEFILYYAPKL 299  
Db 270 SNOTHIEIILYKSCDLTTLILHNLKRGVGTLLVEAGPTTTSFSEIYYIDEFILYYAPKL 329  
QY 300 GSGSNYQFYQTNVDVIEIPDANQFEIVHSELNQNVLTLRKK 342  
Db 330 GSGSNYQFYQTNVDVIEIPDANQFEIVHSELNQNVLTLRKK 372

## RESULT 3

US-09-611-529-5797  
; Sequence 5797 Application US/09611529  
; GENERAL INFORMATION:

; APPLICANT: George H. Shimer, Jr.  
; APPLICANT: George H. Miller  
; APPLICANT: Roberta S. Hare  
; APPLICANT: Karen J. Shaw  
; TITLE OF INVENTION: Staphylococcus aureus Related Compositions and Methods  
; FILE REFERENCE: 1034/1C963USI

; CURRENT APPLICATION NUMBER: US/09/611,529  
; CURRENT FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/417,811  
; PRIOR FILING DATE: 1999-10-14  
; PRIOR APPLICATION NUMBER: US 09/353,718  
; PRIOR FILING DATE: 1999-07-14  
; PRIOR APPLICATION NUMBER: US 09/266,557  
; PRIOR FILING DATE: 1999-03-11  
; PRIOR APPLICATION NUMBER: US 09/266,556  
; PRIOR FILING DATE: 1999-03-11  
; PRIOR APPLICATION NUMBER: US 09/266,555  
; PRIOR FILING DATE: 1999-03-11  
; PRIOR APPLICATION NUMBER: US 09/266,542  
; PRIOR FILING DATE: 1999-03-11  
; PRIOR APPLICATION NUMBER: US 09/266,541  
; PRIOR FILING DATE: 1999-03-11  
; PRIOR APPLICATION NUMBER: US 09/037,934  
; PRIOR FILING DATE: 1998-03-10  
; PRIOR APPLICATION NUMBER: US 09/036,720  
; PRIOR FILING DATE: 1998-03-06  
; PRIOR APPLICATION NUMBER: US 09/036,338  
; PRIOR FILING DATE: 1998-03-06  
; PRIOR APPLICATION NUMBER: US 09/036,334  
; PRIOR FILING DATE: 1998-03-06  
; PRIOR APPLICATION NUMBER: US 09/036,221  
; PRIOR FILING DATE: 1998-03-06  
; PRIOR APPLICATION NUMBER: US 09/036,137  
; PRIOR FILING DATE: 1998-03-06  
; PRIOR APPLICATION NUMBER: US 09/036,082  
; PRIOR FILING DATE: 1998-03-06  
; PRIOR APPLICATION NUMBER: US 09/036,081  
; PRIOR FILING DATE: 1998-03-06  
; PRIOR APPLICATION NUMBER: US 09/036,079  
; PRIOR FILING DATE: 1998-03-06  
; PRIOR APPLICATION NUMBER: US 09/035,913  
; PRIOR FILING DATE: 1998-03-06  
; PRIOR APPLICATION NUMBER: US 09/035,744  
; PRIOR FILING DATE: 1998-03-06  
; PRIOR APPLICATION NUMBER: US 08/827,356

```
; PRIOR FILING DATE: 1997-04-01
; PRIOR APPLICATION NUMBER: US 08/831,156
; PRIOR FILING DATE: 1997-04-01
; PRIOR APPLICATION NUMBER: US 60/014,477
; PRIOR FILING DATE: 1996-04-01
; PRIOR APPLICATION NUMBER: US 60/016,743
; PRIOR FILING DATE: 1996-05-02
; PRIOR APPLICATION NUMBER: US 60/020,016
; PRIOR FILING DATE: 1996-06-14
; NUMBER OF SEQ ID NOS: 7451
; SEQ ID NO 5797
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-611-529-5797

Query Match      98.4%; Score 1759.5; DB 20; Length 372;
Best Local Similarity 98.8%; Pred. No. 2.4e-174;
Matches 339; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 MDYAIQANNVGQOTGPNPPVGVAVVYVNEGRIVGIGAHLRKGDHAEVQALDMAQQNAEGA 60
Db 30 MDYAIQANNVGQOTGPNPPVGVAVVYVNEGRIVGIGAHLRKGDHAEVQALDMAQQNAEGA 89
QY 61 TIITILEPCSHFGSTPPCVNKKIIDCKIAKVYATKDNSLDTHGDETLRAHGIEVECVDD 120
Db 90 TIITILEPCSHFGSTPPCVNKKIIDCKIAKVYATKDNSLDTHGDETLRAHGIEVECVDD 149
QY 121 RASQLYQDFEFKAKAKOLPOITVKVVSASLDGKQANDNGQOWITNKEVKQDVYKLRHDA 180
Db 150 RASQLYQDFEFKAKAKOLPOITVKVVSASLDGKQANDNGQOWITNKEVKQDVYKLRHDA 209
QY 181 VLTGRRTVELDDPYTTRIQDGKNPIKIVLSKSGNIHFNOQIYODESTPIWIYENPNLT 240
Db 210 VLTGRRTVELDDPYTTRIQDGKNPIKIVLSKSGNIHFNOQIYODESTPIWIYENPNLT 269
QY 241 SNOTHEIILYKSCDLTTILHNLKRGVGTLLVEAGPTTISEF-SIYYIDEFILYAPKL 299
Db 270 SNOTHEIILYKSCDLTTILHNLKRGVGTLLVEAGPTTISEFLOSNYIDEFILYAPKL 329
QY 300 IGGSGNYQFYQTNVDIEIPDANQFEIVHSELLNQNVKLTLRKK 342
Db 330 IGGSGNYQFYQTNVDIEIPDANQFEIVHSELLNQNVKLTLRKK 372

RESULT 4
US-09-450-969-6115
; Sequence 6115, Application US/09450969
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: PATH99-09A
; CURRENT APPLICATION NUMBER: US/09/450,969
; CURRENT FILING DATE: 1999-11-29
; NUMBER OF SEQ ID NOS: 7544
; SEQ ID NO 6115
; LENGTH: 350
; TYPE: PRT
; ORGANISM: S.epidermidis
US-09-450-969-6115

Query Match      66.7%; Score 1193.5; DB 18; Length 350;
Best Local Similarity 65.0%; Pred. No. 1.9e-115;
Matches 223; Conservative 61; Mismatches 58; Indels 1; Gaps 1;

QY 1 MDYAIQANNVGQOTGPNPPVGVAVVYVNEGRIVGIGAHLRKGDHAEVQALDMAQQNAEGA 60
Db 8 MDAIQLAKMNVGQOTGPNPPVGVAVVYVNEGRIVGLGHLKGDHAEVQATIEAGLNTQGA 67
QY 61 TIITILEPCSHFGSTPPCVNKKIIDCKIAKVYATKDNSLDTHGDETLRAHGIEVECVDD 120
```

```
Db 68 TIYVLEPCTHHGSTPPCVDKIIIEAGISKVIYAVKDTTLVSKGDEILREAGIEVEFYNE 127
QY 121 RASQLYQDFEFKAKAKOLPOITVKVVSASLDGKQANDNGQOWITNKEVKQDVYKLRHDA 180
Db 128 NAAALYRDEFTAKRNEVPEVTVKVSLSLGGKQATDNESKWIINKEVKEDVYQLRHDA 187
QY 181 VLTGRRTVELDDPYTTRIQDGKNPIKIVLSKSGNIHFNOQIYODESTPIWIYENPNLT 240
Db 188 VITGRRTIEADNPLYTRVPDGHPIRVILSKGQDFNQOIFKDTASEIWIYENKELK 247
QY 241 SNOTHEIILYKSCDLTTILHNLKRGVGTLLVEAGPTTISEF-SIYYIDEFILYAPKL 299
Db 248 TNKSFILINISNCDTTLQDLYQRGKLLVEAGPNITTSQFLQSKHLNELIYIAPKL 307
QY 300 IGGSGNYQFYQTNVDIEIPDANQFEIVHSELLNQNVKLTLRKK 342
Db 308 IGGSGKHQYKTDVIDLPEATQFEIVDSKLINQNLKLRKK 350

RESULT 5
US-60-215-161-5966
; Sequence 5966, Application US/60215161
; GENERAL INFORMATION:
; APPLICANT: Corbin, David R.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Huesing, Joseph E.
; APPLICANT: Krasomil-Osterfeld, Karina C.
; APPLICANT: Malvar, Thomas M.
; APPLICANT: Slater, Steven C.
; APPLICANT: Spiridonov, Sergei
; APPLICANT: Hinkle, Gregory J.
; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51847)A
; CURRENT APPLICATION NUMBER: US/60/215,161
; CURRENT FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 8409
; SEQ ID NO 5966
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Xenorhabdus sp.
US-60-215-161-5966

Query Match      27.9%; Score 500; DB 27; Length 379;
Best Local Similarity 34.4%; Pred. No. 3.7e-43;
Matches 128; Conservative 67; Mismatches 143; Indels 34; Gaps 9;

QY 1 MDYAIQANNVGQOTGPNPPVGVAVVYVNEGRIVGIGAHLRKGDHAEVQALDMAQQNAEGA 60
Db 8 MSRALELAYOGRFTTSPNPNVGCIVVKDQIVGEGPHLRAGEPHAEVHALRMAGERAKGA 67
QY 61 TIITILEPCSHFGSTPPCVNKKIIDCKIAKVYATKDNS--LDTHGDETLRAHGIEVE-CV 117
Db 68 TAYVTLEPCSHHGKTTPCSDALIAAGISRVVAMQDPNQVAGRGYKLYQQAGISYEHGV 127
QY 118 DDERASQLYQDFEFKAKAKOLPOITVKVVSASLDGKQANDNGQOWITNKEVKQDVYKLRH 177
Db 128 MMEQTEMLNKGFLKRMWTGFPYQLQKLGASLDGRTALASGESKWIITSPARQDVQKLRQ 187
QY 178 HDAVLTGRRTVELDDPYTTRIQD-----GKNPIKIVLSKSGNIHFNOQIYQ 224
Db 188 CSAILLSSTATVLADDPSLVVRNNELDATQAVVYPOEALRQPVRIIVDSKNRITPQHVVQ 247
QY 225 DESTPIWI-YTE-----NPNLTSTQTHIEIYIY-----KSCDLTTLHNLKRGV 269
Db 248 -QTGCQWLHAHTDLGSDSDANSADKNEQWPDNIEKILLPVHGTGVDVLLIMQLGROVN 306
QY 270 TLLVEAGPTTISEF-SIYYIDEFILYAPKLIGGSGNYQFYQTNVDIEIPDANQFEIVHS 328
Db 307 SVWAECPFLACALLSLGLVDILYIAPKVLGNSAR-GLFAIPELOKUSDAPFTLIDV 365
QY 329 ELLNQNVKLTLR 340
```





Db 241 YIATSSDKNKMKLYQNHGCEILSIKKGNHIDLSLMLQHLGNMQIDSLVLEGGSLMWS 300  
Qy 282 EFSIYYIDEFLIYAPKLIGGS-----GNQFYQTDNDVIEIPDANQFEIVHSELLNQNV 335  
Db 301 ALEQOIVDELKIYIAPKIFGGSAKFPVGGEGISLPNDAIRLKPYAFSQIGNDYLIESEV 359

RESULT 9  
US-08-979-616-2  
; Sequence 2, Application US/08979616  
; GENERAL INFORMATION:  
; APPLICANT: Palmer, Leslie M.  
; TITLE OF INVENTION: Novel ribg  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dechert Price & Rhoads  
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre  
; CITY: Philadelphia  
; STATE: PA  
; COUNTRY: US  
; ZIP: 19103  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/979,616  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Dickinson, Todd Q  
; REGISTRATION NUMBER: 28,354  
; REFERENCE/DOCKET NUMBER: P50533-8  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 215-994-2252  
; TELEFAX: 215-994-2222  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 366 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-979-616-2

Query Match 26.1%; Score 467; DB 14; Length 366;  
Best Local Similarity 33.1%; Pred. No. 9.5e-40;  
Matches 119; Conservative 65; Mismatches 145; Indels 30; Gaps 9;  
Qy 1 MDYAIQIANMVGGOTGVNPPVGVVNEGRIVGTGAHLRKDGKHAEOALDMAQOAEGA 60  
Db 7 MKAIAKLAQAGAVNPNMVGAIIVDNHITIGQYHEFFGGPHAEARNALKNCRKSPVGA 66  
Qy 61 TIYITLPCSHFGSTPPCVNKKIIDCKIAKVYATKNSLDTHGDETLRAHGLEV---- 114  
Db 67 TLYVTLEPCCHFGTKPTCIDAIIDSGITRVVIGSLDCNPVSGKGVKILENNIQQVTGVI 126  
Qy 115 ---ECVDDERASQLYQDFFAKAKOLPOITVKVSASLDGKOANDNGOSOWITNKVKODV 171  
Db 127 LENECNL-----LIKFRKYITHQVYPVPMKYAMSDGKIATKTNSKWIIEEARKHV 180  
Qy 172 YKLRRHDAVLGTRRTVELDDPQVTTTRIQDGKNPIKVLKSGNIHFNQOIQY---DEST 228  
Db 181 HOLRHYVSAINVGVTVIQDDPLLTCLRECKNPIRICDHLRTPLTSKIVKTANDIKT 240  
Qy 229 PIWIYENPN-LTNSQTH-TEIYILKS-----CDLTTLHNLKKGVTLLVEAGPTT-S 281  
Db 241 YIATSSDKNKMKLYQNHGCEILSIKKGNHIDLSLMLQHLGNMQIDSLVLEGGSLMWS 300

Qy 282 EFSIYYIDEFLIYAPKLIGGS-----GNQFYQTDNDVIEIPDANQFEIVHSELLNQNV 335  
Db 301 ALEQOIVDELKIYIAPKIFGGSAKFPVGGEGISLPNDAIRLKPYAFSQIGNDYLIESEV 359  
RESULT 10  
US-09-252-691-7588  
; Sequence 7588, Application US/09252691B  
; GENERAL INFORMATION:  
; APPLICANT: Keith G. Weinstock et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROBACTE  
; TITLE OF INVENTION: CLOACAE FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.135  
; CURRENT APPLICATION NUMBER: US/09/252,691B  
; CURRENT FILING DATE: 1999-02-18  
; NUMBER OF SEQ ID NOS: 11324  
; SEQ ID NO 7588  
; LENGTH: 402  
; TYPE: PRT  
; ORGANISM: Enterobacter cloacae  
US-09-252-691-7588

Query Match 25.9%; Score 464; DB 17; Length 402;  
Best Local Similarity 33.7%; Pred. No. 2.3e-39;  
Matches 115; Conservative 52; Mismatches 128; Indels 46; Gaps 7;  
Qy 1 MDYAIQIANMVGGOTGVNPPVGVVNEGRIVGTGAHLRKDGKHAEOALDMAQOAEGA 60  
Db 42 MARAMKLAQGRGRTTHPNPNVGCIVKDGIEIVGEGFYRAGEPHAEVHALRMAGEKARGA 101  
Qy 61 TIYITLPCSHFGSTPPCVNKKIIDCKIAKVYATKNSLDTHGDETLR-----AHGI 112  
Db 102 TAYVTLEPCSHHGRTPPCCREALIAGVSRVVAAMQDPNPQVAGRGLYRLQOEGIDVSHGL 161  
Qy 113 EVECVDDERASQLYQDFFAKAKOLPOITVKVSASLDGKOANDNGOSOWITNKVKODVY 172  
Db 162 MM-----QDAELNKGFLKRMRTGFPFIQLKLGASLDGRTAMANGESOWITSPQARRDVQ 216  
Qy 173 KLRHRDAVLGTRRTVELDDPQVTTTRIQD-----GKNPIKVLKSGNIHFN 219  
Db 217 RLRAQSHAILTSETVILADDPAMTVRWEELNADTQALYPOENLRQPLRIIDSQNRVYPE 276  
Qy 220 QOIYQD-----ESTPIWIY-TENPNLT-SNOTHEIYILKSCDLTTLHNLKRGV 268  
Db 277 HRIVQOPGETWIARTKEDTREWPGQVRSITVPEHNGHL-----DLVVLMLLKGQOV 328  
Qy 269 GTLLVEAGPTTTSF-FSIYYIDEFLIYAPKLIGSGNYQF 308  
Db 329 NSIWEAGPTLAGALLQAGLVDELIVVAPKLLGNDARGLF 369

RESULT 11  
PCT-US97-02318-354  
; Sequence 354, Application PC/TUS9702318  
; GENERAL INFORMATION:  
; APPLICANT: Black, Michael  
; APPLICANT: Burnham, Martin  
; APPLICANT: Hodgson, John  
; APPLICANT: Knowles, David  
; APPLICANT: Nicholas, Richard  
; APPLICANT: Pratt, Julie  
; APPLICANT: Reichard, Raymond  
; APPLICANT: Rosenberg, Martin  
; APPLICANT: Ward, Judith  
; TITLE OF INVENTION: Novel Compounds  
; NUMBER OF SEQUENCES: 1166  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Smithkline Beecham Corporation  
; STREET: 709 Swedeland Road  
; CITY: King of Prussia  
; STATE: PA





**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 14, 2000, 10:09:29 ; Search time 72.12 Seconds  
(without alignments)  
185.968 Million cell updates/sec

Title: US-08-978-456-4  
Perfect score: 498  
Sequence: 1 MDYALQLPMVQGTGVNPP.....NKIIDCKTAXVVLXNXXRFR 99

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 866916 seqs, 135474527 residues

Total number of hits satisfying chosen parameters: 866916

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_AA:\*

1:	/cgn2_6/ptodata/2/paa/PCTUS_COMB.pep.*
2:	/cgn2_6/ptodata/2/paa/US06_COMB.pep.*
3:	/cgn2_6/ptodata/2/paa/US07_COMB.pep.*
4:	/cgn2_6/ptodata/2/paa/US080_COMB.pep.*
5:	/cgn2_6/ptodata/2/paa/US081_COMB.pep.*
6:	/cgn2_6/ptodata/2/paa/US082_COMB.pep.*
7:	/cgn2_6/ptodata/2/paa/US083_COMB.pep.*
8:	/cgn2_6/ptodata/2/paa/US084A_COMB.pep.*
9:	/cgn2_6/ptodata/2/paa/US084B_COMB.pep.*
10:	/cgn2_6/ptodata/2/paa/US085_COMB.pep.*
11:	/cgn2_6/ptodata/2/paa/US086_COMB.pep.*
12:	/cgn2_6/ptodata/2/paa/US087_COMB.pep.*
13:	/cgn2_6/ptodata/2/paa/US088_COMB.pep.*
14:	/cgn2_6/ptodata/2/paa/US089_COMB.pep.*
15:	/cgn2_6/ptodata/2/paa/US090_COMB.pep.*
16:	/cgn2_6/ptodata/2/paa/US091_COMB.pep.*
17:	/cgn2_6/ptodata/2/paa/US092_COMB.pep.*
18:	/cgn2_6/ptodata/2/paa/US093_COMB.pep.*
19:	/cgn2_6/ptodata/2/paa/US094_COMB.pep.*
20:	/cgn2_6/ptodata/2/paa/US095_COMB.pep.*
21:	/cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep.*
22:	/cgn2_6/ptodata/2/paa/US05_NEW_COMB.pep.*
23:	/cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep.*
24:	/cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep.*
25:	/cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep.*
26:	/cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep.*
27:	/cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	488	98.0	99	1	PCT-US97-02318-354 Sequence 354, App
2	488	98.0	99	1	PCT-US98-24857-4 Sequence 4, Appli
3	488	98.0	99	14	US-08-903-470-354 Sequence 354, App
4	451	90.6	342	1	PCT-US98-24857-2 Sequence 2, Appli
5	451	90.6	372	13	US-08-827-356-4163 Sequence 4163, Ap
6	451	90.6	372	20	US-09-611-529-5797 Sequence 5797, Ap

Sequence 6115, Ap  
Sequence 5966, Ap  
Sequence 7588, Ap  
Sequence 78, Appli  
Sequence 10, Appli  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 4805, Ap  
Sequence 14, Appli  
Sequence 5462, Ap  
Sequence 5675, Ap  
Sequence 5, Appli  
Sequence 873, App  
Sequence 3272, Ap  
Sequence 4795, Ap  
Sequence 15, Appli  
Sequence 5704, Ap  
Sequence 7583, Ap  
Sequence 4901, Ap  
Sequence 4436, Ap  
Sequence 650, App  
Sequence 650, App  
Sequence 15, Appli  
Sequence 5242, Ap  
Sequence 3411, Ap  
Sequence 4977, Ap  
Sequence 6453, Ap  
Sequence 1112, Ap  
Sequence 378, App  
Sequence 6593, Ap  
Sequence 8132, Ap  
Sequence 8133, Ap  
Sequence 8134, Ap  
Sequence 7636, Ap  
Sequence 4, Appli  
Sequence 16, Appli  
Sequence 8, Appli

US-09-450-969-6115  
US-60-215-161-5966  
US-09-252-691-7588  
US-09-614-912-78  
US-60-170-906-10  
PCT-US98-25010-2  
US-08-979-616-2  
US-09-583-110-4805  
US-60-170-906-14  
US-09-328-352-5462  
US-09-107-532-5675  
US-09-369-955-5  
US-09-438-185-873  
US-09-540-236-3272  
US-60-128-476-4795  
US-09-064-935-15  
US-60-215-161-5704  
US-09-328-352-7583  
US-09-107-532-4012  
US-09-134-000-4901  
US-09-583-110-4436  
PCT-US97-14436-650  
US-08-911-503-650  
US-08-911-503A-650  
US-60-170-906-15  
US-09-450-969-5242  
US-08-827-356-3411  
US-09-611-529-4977  
US-09-252-691-6455  
PCT-US98-06371-1112  
US-08-902-615A-378  
US-09-328-352-6593  
US-08-993-002A-8132  
US-08-993-002A-8133  
US-08-993-002A-8134  
US-60-215-161-7636  
US-08-403-768-4  
PCT-US99-10985-16  
US-07-928-506-8

ALIGNMENTS

RESULT 1  
PCT-US97-02318-354  
; Sequence 354, Application PC/TUS9702318  
; GENERAL INFORMATION:  
; APPLICANT: Black, Michael  
; APPLICANT: Burnham, Martin  
; APPLICANT: Hodgson, John  
; APPLICANT: Knowles, David  
; APPLICANT: Nicholas, Richard  
; APPLICANT: Pratt, Julie  
; APPLICANT: Reichard, Raymond  
; APPLICANT: Rosenberg, Martin  
; APPLICANT: Ward, Judith  
; TITLE OF INVENTION: Novel Compounds  
; NUMBER OF SEQUENCES: 1166  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SmithKline Beecham Corporation  
; STREET: 709 Swedeland Road  
; CITY: King of Prussia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19406-0939  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US97/02318

```

; FILING DATE: 19-FEB-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/011,888
; FILING DATE: 20-FEB-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmli, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: GM50006
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
;
; INFORMATION FOR SEQ ID NO: 354:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 99 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US97-02318-354

Query Match 98.0%; Score 488; DB 1; Length 99;
Best Local Similarity 100.0%; Pred. No. 7.4e-56;
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MDYATQLPNMVGXTCVNPVPGAVVNVNIGRIGVIGAHLRKGDHAEVQALDMAQXNAEGA 60
Db 1 MDYATQLPNMVGXTCVNPVPGAVVNVNIGRIGVIGAHLRKGDHAEVQALDMAQXNAEGA 60
QY 61 TIYITLPCSHFGSTPPCVNKKIIDCKIAAXVLXNXRQFR 99
Db 61 TIYITLPCSHFGSTPPCVNKKIIDCKIAAXVLXNXRQFR 99

```

```

RESULT 2
PCT-US98-24857-4
; Sequence 4, Application PC/TUS9824857A
; GENERAL INFORMATION:
; TITLE OF INVENTION: Novel ribG
; FILE REFERENCE: P50444-9
; CURRENT APPLICATION NUMBER: PCT/US98/24857A
; CURRENT FILING DATE: 1998-11-23
; EARLIER APPLICATION NUMBER: 08/978,456
; EARLIER FILING DATE: 1997-11-25
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(99)
; OTHER INFORMATION: Xaa = Any Amino Acid
; PCT-US98-24857-4

```

```

Query Match 98.0%; Score 488; DB 1; Length 99;
Best Local Similarity 100.0%; Pred. No. 7.4e-56;
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MDYATQLPNMVGXTCVNPVPGAVVNVNIGRIGVIGAHLRKGDHAEVQALDMAQXNAEGA 60
Db 1 MDYATQLPNMVGXTCVNPVPGAVVNVNIGRIGVIGAHLRKGDHAEVQALDMAQXNAEGA 60
QY 61 TIYITLPCSHFGSTPPCVNKKIIDCKIAAXVLXNXRQFR 99
Db 61 TIYITLPCSHFGSTPPCVNKKIIDCKIAAXVLXNXRQFR 99

```

```

RESULT 3
US-08-903-470-354
; Sequence 354, Application US/08903470
; GENERAL INFORMATION:
; APPLICANT: Black, Michael
; APPLICANT: Burnham, Martin
; APPLICANT: Hodgson, John
; APPLICANT: Knowles, David
; APPLICANT: Nicholas, Richard
; APPLICANT: Pratt, Julie
; APPLICANT: Reichard, Raymond
; APPLICANT: Rosenberg, Martin
; APPLICANT: Ward, Judith
; TITLE OF INVENTION: Novel Compounds
; NUMBER OF SEQUENCES: 1166
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/903,470
; FILING DATE: 24-JUL-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/011,888
; FILING DATE: 20-FEB-1996
; APPLICATION NUMBER: PCT/US97/02318
; FILING DATE: 19-FEB-1997
; APPLICATION NUMBER: PCT/US97/02547
; FILING DATE: 19-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmli, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: GM50006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 354:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 99 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-903-470-354

```

```

Query Match 98.0%; Score 488; DB 14; Length 99;
Best Local Similarity 100.0%; Pred. No. 7.4e-56;
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MDYATQLPNMVGXTCVNPVPGAVVNVNIGRIGVIGAHLRKGDHAEVQALDMAQXNAEGA 60
Db 1 MDYATQLPNMVGXTCVNPVPGAVVNVNIGRIGVIGAHLRKGDHAEVQALDMAQXNAEGA 60
QY 61 TIYITLPCSHFGSTPPCVNKKIIDCKIAAXVLXNXRQFR 99
Db 61 TIYITLPCSHFGSTPPCVNKKIIDCKIAAXVLXNXRQFR 99

```

```

RESULT 4
PCT-US98-24857-2
; Sequence 2, Application PC/TUS9824857A
; GENERAL INFORMATION:

```

```
; APPLICANT: SmithKline Beecham Corporation
; TITLE OF INVENTION: Novel ribg
; FILE REFERENCE: P50444-9
; CURRENT APPLICATION NUMBER: PCT/US98/24857A
; CURRENT FILING DATE: 1998-11-23
; EARLIER APPLICATION NUMBER: 08/978,456
; EARLIER FILING DATE: 1997-11-25
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 342
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
PCT-US98-24857-2

Query Match          90.6%; Score 451; DB 1; Length 342;
Best Local Similarity 95.6%; Pred. No. 2.6e-50;
Matches 87; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MDYAIQLPMVQGTGVNPPVGVVNVNIGRIGVGAHLRKGDKHAEVQALDMAQXNAEGA 60
Db 1 MDYAIQLANMVQGTGVNPPVGVVNVNIGRIGVGAHLRKGDKHAEVQALDMAQXNAEGA 60

Qy 61 TIYITLPCSHFGSTPPCVNKKIIDCKIAVV 91
Db 61 TIYITLPCSHFGSTPPCVNKKIIDCKIAKV 91

RESULT 5
US-08-827-356-4163
; Sequence 4163, Application US/08827356
; GENERAL INFORMATION:
; APPLICANT: George H. Shimer, Jr.
; APPLICANT: George H. Miller
; APPLICANT: Roberta S. Hare
; APPLICANT: Karen J. Shaw
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS RELATED
; TITLE OF INVENTION: COMPOSITIONS AND METHODS
; NUMBER OF SEQUENCES: 5574
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schering-Plough Corporation
; STREET: 2000 Galloping Hill Road
; CITY: Kenilworth
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07033-0530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/827,356
; FILING DATE: 01-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/014,477
; FILING DATE: 01-APR-1996
; APPLICATION NUMBER: 60/016,743
; FILING DATE: 02-MAY-1996
; APPLICATION NUMBER: 60/020,016
; FILING DATE: 14-JUN-1996
; INFORMATION FOR SEQ ID NO: 4163:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 372 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: misc_feature
;
```

```
; LOCATION: 1...372
US-08-827-356-4163

Query Match          90.6%; Score 451; DB 13; Length 372;
Best Local Similarity 95.6%; Pred. No. 2.9e-50;
Matches 87; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MDYAIQLPMVQGTGVNPPVGVVNVNIGRIGVGAHLRKGDKHAEVQALDMAQXNAEGA 60
Db 30 MDYAIQLANMVQGTGVNPPVGVVNVNIGRIGVGAHLRKGDKHAEVQALDMAQXNAEGA 89

Qy 61 TIYITLPCSHFGSTPPCVNKKIIDCKIAVV 91
Db 90 TIYITLPCSHFGSTPPCVNKKIIDCKIAKV 120

RESULT 6
US-09-611-529-5797
; Sequence 5797, Application US/09611529
; GENERAL INFORMATION:
; APPLICANT: George H. Shimer, Jr.
; APPLICANT: George H. Miller
; APPLICANT: Roberta S. Hare
; APPLICANT: Karen J. Shaw
; TITLE OF INVENTION: Staphylococcus aureus Related Compositions and Methods
; FILE REFERENCE: 1034/1C963US1
; CURRENT APPLICATION NUMBER: US/09/611,529
; CURRENT FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/417,811
; PRIOR FILING DATE: 1999-10-14
; PRIOR APPLICATION NUMBER: US 09/353,718
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: US 09/266,557
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,556
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,555
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,542
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,541
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/037,934
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: US 09/036,720
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/036,338
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/036,334
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/036,221
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/036,137
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/036,082
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/036,081
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/036,079
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/035,913
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/035,744
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 08/827,356
; PRIOR FILING DATE: 1997-04-01
; PRIOR APPLICATION NUMBER: US 08/831,156
; PRIOR FILING DATE: 1997-04-01
; PRIOR APPLICATION NUMBER: US 60/014,477
; PRIOR FILING DATE: 1996-04-01
; PRIOR APPLICATION NUMBER: US 60/016,743
; PRIOR FILING DATE: 1996-05-02
;
```

```
; PRIOR APPLICATION NUMBER: US 60/020,016
; PRIOR FILING DATE: 1996-06-14
; NUMBER OF SEQ ID NOS: 7451
; SEQ ID NO 5797
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-611-529-5797

Query Match          90.6%; Score 451; DB 20; Length 372;
Best Local Similarity 95.6%; Pred. No. 2.9e-50;
Matches 87; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 MDVAIQLPNMVQGTGYNPPVAVVNEGRIVGIGAHLRKGDHAEVQALDMAQXNAEGA 60
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 30 MDVAIQLANVQGTGYNPPVAVVNEGRIVGIGAHLRKGDHAEVQALDMAQXNAEGA 89

QY 61 TIYTILEPCSHFGSTPPCVNKIIDCKIAXVV 91
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 90 TIYTILEPCSHFGSTPPCVNKIIDCKIAKVV 120

RESULT 7
US-09-450-969-6115
; Sequence 6115, Application US/09450969
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: 107196.135
; CURRENT FILING DATE: 1999-11-29
; NUMBER OF SEQ ID NOS: 7544
; SEQ ID NO 6115
; LENGTH: 350
; TYPE: PRT
; ORGANISM: S.epidermidis
US-09-450-969-6115

Query Match          69.7%; Score 347; DB 18; Length 350;
Best Local Similarity 70.3%; Pred. No. 1e-36;
Matches 64; Conservative 13; Mismatches 14; Indels 0; Gaps 0;

QY 1 MDVAIQLPNMVQGTGYNPPVAVVNEGRIVGIGAHLRKGDHAEVQALDMAQXNAEGA 60
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 8 MDVAIQAKVNGOTGYNPPVGVVKNRIVGLGAHLKKGDKHAEVQALDMAQXNAEGA 67

QY 61 TIYTILEPCSHFGSTPPCVNKIIDCKIAXVV 91
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 68 TIVVLEPCTHHGSTPPCVNKIIDCKIAXVV 98

RESULT 8
US-60-215-161-5966
; Sequence 5966, Application US/60215161
; GENERAL INFORMATION:
; APPLICANT: Corbin, David R.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Huesing, Joseph E.
; APPLICANT: Krasomil-Osterfeld, Karina C.
; APPLICANT: Malvar, Thomas M.
; APPLICANT: Slater, Steven C.
; APPLICANT: Spiridonov, Sergei
; APPLICANT: Hinkle, Gregory J.
; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51847)A
; CURRENT APPLICATION NUMBER: US/60/215,161
; CURRENT FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 8409
; SEQ ID NO 5966
; LENGTH: 379
```

```
; TYPE: PRT
; ORGANISM: Xenorhabdus sp.
US-60-215-161-5966

Query Match          46.6%; Score 232; DB 27; Length 379;
Best Local Similarity 51.1%; Pred. No. 1.2e-21;
Matches 47; Conservative 13; Mismatches 32; Indels 0; Gaps 0;

QY 1 MDVAIQLPNMVQGTGYNPPVAVVNEGRIVGIGAHLRKGDHAEVQALDMAQXNAEGA 60
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 8 MSRALELAYOGRFTTTPNPNVGCIVVKDQIVGEGHFRAGEPHAEVHALRMAGERAKGA 67

QY 61 TIYTILEPCSHFGSTPPCVNKIIDCKIAXVV 92
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 68 TAVVTLEPCSHHGKTPPCSDALIAAGISRVVV 99

RESULT 9
US-09-252-691-7588
; Sequence 7588, Application US/09252691B
; GENERAL INFORMATION:
; APPLICANT: Keith G. Weinstock et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROBACTE
; FILE REFERENCE: 107196.135
; CURRENT FILING DATE: 1999-02-18
; NUMBER OF SEQ ID NOS: 11324
; SEQ ID NO 7588
; LENGTH: 402
; TYPE: PRT
; ORGANISM: Enterobacter cloacae
US-09-252-691-7588

Query Match          45.0%; Score 224; DB 17; Length 402;
Best Local Similarity 50.5%; Pred. No. 1.5e-20;
Matches 46; Conservative 10; Mismatches 35; Indels 0; Gaps 0;

QY 1 MDVAIQLPNMVQGTGYNPPVAVVNEGRIVGIGAHLRKGDHAEVQALDMAQXNAEGA 60
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 42 MARAKMLAQRFTTTPNPNVGCIVVKDQIVGEGHFRAGEPHAEVHALRMAGERAKGA 101

QY 61 TIYTILEPCSHFGSTPPCVNKIIDCKIAXVV 91
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 102 TAVVTLEPCSHHGKTPPCCEALIAAGVSRVV 132

RESULT 10
US-09-614-912-78
; Sequence 78, Application US/09614912
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Rafalski, Antoni
; APPLICANT: Orozco, Buddy
; APPLICANT: Miao, Gou-Hau
; APPLICANT: Famodu, Omolayo O.
; APPLICANT: Lee, Jian Ming
; APPLICANT: Sakai, Hajime
; APPLICANT: Weng, Zude
; APPLICANT: Caimi, Perry G
; APPLICANT: Anderson, Shawn
; TITLE OF INVENTION: Plant Metabolism Genes
; FILE REFERENCE: BB1378 US NA
; CURRENT APPLICATION NUMBER: US/09/614,912
; CURRENT FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: 60/143,401
; PRIOR FILING DATE: 1999-07-12
; PRIOR APPLICATION NUMBER: 60/143,412
; PRIOR FILING DATE: 1999-07-12
; PRIOR APPLICATION NUMBER: 60/146,650
; PRIOR FILING DATE: 1999-07-30
```



```

; ATTORNEY/AGENT INFORMATION:
; NAME: Dickinson, Todd Q
; REGISTRATION NUMBER: 28,354
; REFERENCE/DOCKET NUMBER: P50533-8
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-994-2252
; TELEFAX: 215-994-2222
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 366 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-08-979-616-2

```



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 14, 2000, 10:06:42 ; Search time 28.37 Seconds  
(without alignments)  
412.206 Million cell updates/sec

Title: US-08-978-456-2

Perfect score: 1789

Sequence: 1 MDYAIQLANMVGQTGVNPP.....FEIVHSLLNONVKLTLRKK 342

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_36.\*  
1: /SIDSL/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /SIDSL/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /SIDSL/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /SIDSL/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /SIDSL/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /SIDSL/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /SIDSL/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /SIDSL/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /SIDSL/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /SIDSL/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /SIDSL/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /SIDSL/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /SIDSL/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /SIDSL/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /SIDSL/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /SIDSL/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /SIDSL/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /SIDSL/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /SIDSL/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /SIDSL/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /SIDSL/gcgdata/geneseq/geneseq/AA2000.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1789	100.0	342	Y15913	Pyrimidine deaminase
2	547	30.6	361	12 R10070	B. subtilis rib ORF
3	547	30.6	1855	20 Y21803	B. subtilis rib op
4	547	30.6	1855	21 Y83271	Polypeptide encode
5	467	26.1	366	20 Y17949	S. pneumoniae ribg
6	451	25.2	99	18 W28084	Amino acid sequenc
7	451	25.2	99	20 Y15914	Pyrimidine deamina
8	435	24.3	376	20 Y35519	Chlamydia pneumoni
9	393	22.0	410	19 W55031	Actinobacillus ple
10	393	22.0	410	20 Y21972	App ribg gene prod
11	356	19.9	401	21 Y96284	Sinorhizobium meli
12	277	15.5	249	20 W81550	Bacillus subtilis

13	234.5	13.1	261	20	Y37427	Amino acid sequenc
14	194	10.8	170	20	Y37428	Amino acid sequenc
15	153	8.6	155	19	Y86119	S. pneumoniae deri
16	140	7.8	344	19	W98756	H. pylori GHPO 107
17	138.5	7.7	373	20	Y53913	Amino acid sequenc
18	128.5	7.2	158	11	R08285	Thermally stable c
19	128.5	7.2	158	16	R68507	Yeast cytosine-dea
20	127.5	7.1	158	21	Y55104	Chimeric cytosine
21	125.5	7.0	158	21	Y55103	Chimeric cytosine
22	122.5	6.8	158	21	Y55101	Chimeric cytosine
23	121.5	6.8	246	16	R88283	Ashbya gossypii HT
24	119.5	6.7	158	21	Y55102	Chimeric cytosine
25	117.5	6.6	609	16	R88279	Ashbya gossypii DR
26	117	6.5	244	15	R52829	HTP-reductase. Sa
27	111	6.2	163	20	Y35657	Chlamydia pneumoni
28	107.5	6.0	176	20	Y37278	Amino acid sequenc
29	104	5.8	591	15	R52825	DRAP-deaminase. S
30	101.5	5.7	245	18	W89774	Staphylococcus aur
31	98	5.5	1277	21	Y91928	Paenibacillus pabu
32	93	5.2	670	15	R65795	APP751 beta-amylol
33	93	5.2	770	10	P94775	Novel amyloid prec
34	93	5.2	770	13	R26340	APP770. Homo sapi
35	93	5.2	770	14	R41546	Mutated APP770 exo
36	93	5.2	770	15	R63442	Amyloid protein pr
37	93	5.2	770	15	R62505	Amyloid precursor
38	93	5.2	770	18	W19500	APP770 mutant A-be
39	93	5.2	770	18	W19488	APP770 mutant A-be
40	93	5.2	770	18	W19503	APP770 mutant A-be
41	93	5.2	770	18	W19506	APP770 mutant A-be
42	93	5.2	770	18	W19497	APP770 mutant A-be
43	93	5.2	770	18	W19491	APP770 mutant A-be
44	93	5.2	770	18	W19485	APP770 mutant A-be
45	93	5.2	770	18	W19482	APP770 mutant A-be

#### ALIGNMENTS

RESULT 1  
Y15913  
ID Y15913 standard; Protein; 342 AA.  
XX  
AC Y15913;  
XX  
DT 04-AUG-1999 (first entry)  
XX  
DE Pyrimidine deaminase and pyrimidine reductase (ribG) polypeptide.  
XX  
KW Pyrimidine deaminase; pyrimidine reductase; ribG; antagonist;  
KW agonist; antimicrobial; antibiotic; Helicobacter pylori infection;  
KW H. pylori-induced cancer; gastrointestinal carcinoma; gastric ulcer;  
KW gastritis.  
XX  
OS Staphylococcus aureus.  
XX  
PN WO9926475-A1.  
XX  
PD 03-JUN-1999.  
XX  
PF 23-NOV-1998; 98WO-US24857.  
XX  
PR 25-NOV-1997; 97US-0978456.  
XX  
(SMIK ) SMITHKLINE BEECHAM CORP.  
PI Black MT, Burnham MKR, Fedon JC, Hodgson JE, Knowles DJC;  
PI Lonetto MA, Nicholas RO, Palmer LM, Pratt JM, Reichard RW;  
PI Rosenberg M, Trani CM, Ward JM, Warren RL;  
XX WPI; 1999-347572/29.  
DR N-PSDB; X59909.  
XX  
PT New Staphylococcus aureus polypeptide and polynucleotide useful in



```
PR 11-SEP-1990; 90US-0581048.
PR 22-JUN-1989; 89US-0370378.
PR 21-APR-1992; 92US-0873572.
PR 06-FEB-1995; 95US-0384626.
PR 24-AUG-1998; 98US-0138775.
XX
XX (HOFF-) ROCHE VITAMINS INC.
XX
XX Erdenberger T, Hatch RT, Hermann T, Perkins JB;
PI Pero JG, Sloma A;
XX
XX WPI; 1999-418271/35.
DR N-PSDB; X81947.
XX
XX Use of bacterial strains for the over production of riboflavin
PT
XX
XX Example 2; Fig 3A-S; 61pp; English.
XX
XX The invention relates to a method of producing riboflavin that comprises,
CC providing a recombinant bacterium containing an exogenously introduced
CC bacterial ORF (open reading frame) 3 (or ORF5) structural gene and an
CC exogenous promoter, culturing the bacterium, and recovery of the product.
CC The method is useful for the production of large quantities of riboflavin
CC (vitamin B2). Sequences Y21801-803 are protein sequences translated from
CC three different reading frames of B. subtilis rib operon. Y21804-806 are
CC protein sequences translated from three different reading frames of the
CC complementary sequence.
XX
XX Sequence 1855 AA;
SQ
Query Match 30.6%; Score 547; DB 20; Length 1855;
Best Local Similarity 36.1%; Pred. No. 3.6e-42;
Matches 130; Conservative 68; Mismatches 134; Indels 28; Gaps 8;
QY 1 MDYAIQIANVVGQTGYNPPVAVVNEGRIVGICAHLRKGDKEAEVQALDMAQONAEGA 60
Db 372 mklaIdlak9egqdesnplvgavvkdqglvgngahlkgygeahaeVhathmagahaega 431
QY 61 TIYITLPCSHFGSTPCVANKIIDCKIAKVYATKDNS--LDTHGDETLRAHGTEV-ECV 117
Db 432 diyvtlepcshygtppcaeilinsgkrvfamrdpnpiavagrgismkeagievregi 491
QY 118 DDERASOLYQDFFAKAKAQLPQITVKVVSALDQKQANDGQSQWITNKEVKQDYKLRHR 177
Db 492 ladqaerlneklflhmrtgipyvtlkaasldgklatstgdskwitseaarqdaqyrkt 551
QY 178 HDAVLTCRRIVELDDPQYTRIQD-GKNPIKVLKSGNTHFNQOIYQDESTPIWIYT-- 234
Db 552 hqsilvgvtvkadnpsltcrlpnvtkqpvrvliIdtVlsipedakvicdqIaptwifta 611
QY 235 -----ENPNLITSNOTHTETIYVLSKCDLTTLHNLKRGVGLVLAAGPTTTSF-SIYI 288
Db 612 radeekkrksafaynfltieteriqipdvklkllaegimsyveggsavhgsfvkegcf 671
QY 289 DEFILLYAPKLIGG-----SG-NYQFYQTNVDVIEIPDANQFEIVHSELLNQVKLTLR 340
Db 672 qeliifyapkligghapslisgefgsmkdvpIlgtditq-----igrdikItak 723
RESULT 4
Y83271
ID Y83271 standard; Protein; 1855 AA.
XX
AC Y83271;
XX
XX 16-AUG-2000 (first entry)
DT
DE Polypeptide encoded by rib operon of Bacillus subtilis.
XX
KW Riboflavin; rib operon; yeast; bacteria; transformation;
KW transcription element; gene expression; hair loss; skin inflammation;
KW growth disorder; vision; vision; sight.
```

```
XX
OS
PH KEY Location/Qualifiers
FT Misc-difference 5 /note= "Encoded by TAG stop codon"
FT Misc-difference 13 /note= "Encoded by TAA stop codon"
FT Misc-difference 35 /note= "Encoded by TAA stop codon"
FT Misc-difference 72 /note= "Encoded by TAG stop codon"
FT Misc-difference 112 /note= "Encoded by TAG stop codon"
FT Misc-difference 147 /note= "Encoded by TAG stop codon"
FT Misc-difference 149 /note= "Encoded by TGA stop codon"
FT Misc-difference 174 /note= "Encoded by TGA stop codon"
FT Misc-difference 179 /note= "Encoded by TAA stop codon"
FT Misc-difference 206 /note= "Encoded by TGA stop codon"
FT Misc-difference 211 /note= "Encoded by TGA stop codon"
FT Misc-difference 214 /note= "Encoded by TGA stop codon"
FT Misc-difference 235 /note= "Encoded by TAA stop codon"
FT Misc-difference 272 /note= "Encoded by TAA stop codon"
FT Misc-difference 302 /note= "Encoded by TGA stop codon"
FT Misc-difference 311 /label= Encoded by TAA stop codon
FT Misc-difference 314 /note= "Encoded by TGA stop codon"
FT Misc-difference 334 /note= "Encoded by TGA stop codon"
FT Misc-difference 350 /note= "Encoded by TAG stop codon"
FT Misc-difference 359 /note= "Encoded by TAA stop codon"
FT Misc-difference 728 /note= "Encoded by TAA stop codon"
FT Misc-difference 738 /note= "Encoded by TAG stop codon"
FT Misc-difference 783 /note= "Encoded by TAA stop codon"
FT Misc-difference 796 /note= "Encoded by TGA stop codon"
FT Misc-difference 805 /note= "Encoded by TGA stop codon"
FT Misc-difference 852 /note= "Encoded by TGA stop codon"
FT Misc-difference 908 /note= "Encoded by TGA stop codon"
FT Misc-difference 926 /note= "Encoded by TAA stop codon"
FT Misc-difference 929 /note= "Encoded by TAA stop codon"
FT Misc-difference 931 /note= "Encoded by TGA stop codon"
FT Misc-difference 1351 /note= "Encoded by TAA stop codon"
FT Misc-difference 1363 /note= "Encoded by TAA stop codon"
FT Misc-difference 1370 /note= "Encoded by TGA stop codon"
FT Misc-difference 1380 /note= "Encoded by TAG stop codon"
FT Misc-difference 1381 /note= "Encoded by TAG stop codon"
FT
```



CC a sample. ribG can also be used to identify antagonists or agonists.  
CC ribG, or its related nucleic acid, also has use as a vaccine to induce  
CC an immunological response in an animal. Antimicrobial compounds (e.g.  
CC agonists and antagonists of ribG), especially broad-spectrum antibiotics,  
CC may be of use in the treatment of *Helicobacter pylori* infection. This  
CC should decrease the advent of *H. pylori*-induced cancers, such as  
CC gastrointestinal carcinoma. The treatment should also cure gastric ulcers  
CC and gastritis. The present sequence represents the *S. pneumoniae* ribG  
CC polypeptide.  
XX  
SQ Sequence 366 AA;

Query Match 26.1%; Score 467; DB 20; Length 366;  
Best Local Similarity 33.1%; Pred. No. 1e-35;  
Matches 119; Conservative 65; Mismatches 145; Indels 30; Gaps 9;  
QY 1 MDYATOLANVQGTGVNPPGAVVNEGRIVGIGAHLRKDGKHAEOALDMAQONAEGA 60  
DB 7 mklakleqkagayvnpmpvgalivdnhilggyheffgpphaernalkncrkspvga 66  
QY 61 TIYTLEPCSHFGSTPPCVNKKIIDCKIAKVYATKDNS--LDTHGDEFLRAHGIEV--- 114  
DB 67 tlyvtlepcchfgktpccidaiidsgitrsvigslcdnpiyvgkvikleennlqvtvgi 126  
QY 115 ---ECVDDERASQLYQDFFAKAKOLPQITVYKVSASLDGKQANDNGOSQWITNKEVKQDV 171  
DB 127 lenecln-----lksfrkyitghvpyvmkysamsmgdkiatkngskwiteearkhv 180  
QY 172 YKLRHRDAVLTGRVTELDPOVTTTQDCKNPIKVLKSGNIHFNQIYQ---DEST 228  
DB 181 hqlrhysalnmvgvntvldqdpdlctleegknprilicdthlrtpktskivktandikt 240  
QY 229 PIWYVTENPN-LTSNQTH-IEIIVLKS----CDLTTLNLNLYKRGVGLLVEAGPTTT-S 281  
DB 241 yiatssedkmmklyqhngceilsikkgnhidlsslmhqlgmqidslvlegslmws 300  
QY 282 EFSYIIDEFLYAPKLIGGS-----GNVOFYQTNVDVIEIPDANQPEIVHSELLNQNV 335  
DB 301 aleqgvidelkiyapklfggsakfpvggegislndairikpyafsqigndyliesev 359

## RESULT 6

W28084  
ID W28084 standard; Protein; 99 AA.  
XX  
AC W28084;  
XX  
DT 01-SEP-1998 (first entry)  
XX  
DE Amino acid sequence of riboflavin biosynthesis protein RIBG.  
XX  
KW Staphylococcus aureus protein; ribozyme; antisense sequence; control;  
KW Staphylococcal gene; regulatory element; bacterial gene expression;  
KW vaccine; Staphylococcal infection; food poisoning; scaled skin syndrome;  
KW toxic shock syndrome.  
XX  
OS Staphylococcus aureus.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 1..99  
FT /note= "residues designated X are not defined in  
FT the specification"  
XX  
PN W09730070-A1.  
XX  
PD 21-AUG-1997.  
XX  
PF 19-FEB-1997; 97WO-US02318.  
XX  
PR 20-FEB-1996; 96US-0011888.  
XX  
PA (SMIK ) SMITHKLINE BEECHAM CORP.

XX Black MT, Burnham MK, Hodgson JE, Knowles DJC, Nicholas RO;  
PI Pratt JM, Reichard RW, Rosenberg M, Ward JW;  
XX WPI: 1997-424969/39.  
DR N-PSDB; T84023.  
XX Novel polypeptide(s) from *Staphylococcus aureus* strain WCUH29 - used  
PT to isolate antimicrobial compounds, and in vaccines against *S.*  
PT aureus infection  
XX  
PS Claim 6; Page 454-455; 989pp; English.  
XX  
CC The present sequence represents a *Staphylococcus aureus* protein, that,  
CC based on homology with a *Bacillus subtilis* protein, is believed to be a  
CC riboflavin biosynthesis protein RIBG. The DNA sequence was isolated from  
CC a library of clones of *S. aureus* WCUH 29 in *Escherichia coli*. The DNA  
CC sequence can be used in the construction of ribozymes and antisense  
CC sequences to control the expression of staphylococcal genes. The DNA  
CC sequence is also useful as a source of regulatory elements for the  
CC control of bacterial gene expression. The present protein may be used  
CC to produce vaccines to enable a host to produce specific antibodies  
CC with antibacterial action. These vaccines and antibodies would protect  
CC a host against invasion by *S. aureus*, and conditions relating to  
CC Staphylococcal infection, e.g. Staphylococcal food poisoning, scaled  
CC skin syndrome, and toxic shock syndrome.  
XX  
SQ Sequence 99 AA;

Query Match 25.2%; Score 451; DB 18; Length 99;  
Best Local Similarity 95.6%; Pred. No. 4.8e-35;  
Matches 87; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
QY 1 MDYATOLANVQGTGVNPPGAVVNEGRIVGIGAHLRKDGKHAEOALDMAQONAEGA 60  
DB 1 mdyatolpnmvngvntvldqdpdlctleegknprilicdthlrtpktskivktandikt 60  
QY 61 TIYTLEPCSHFGSTPPCVNKKIIDCKIAKV 91  
DB 61 tiytlepcshfgstppcvnkiidckiaxvv 91

## RESULT 7

Y15914  
ID Y15914 standard; Protein; 99 AA.  
XX  
AC Y15914;  
XX  
DT 04-AUG-1999 (first entry)  
XX  
DE Pyrimidine deaminase and pyrimidine reductase (ribG) encoded by ORF.  
XX  
KW Pyrimidine deaminase; pyrimidine reductase; ribG; antagonist;  
KW agonist; antimicrobial; antibiotic; *Helicobacter pylori* infection;  
KW *H. pylori*-induced cancer; gastrointestinal carcinoma; gastric ulcer;  
KW gastritis.  
XX  
OS Staphylococcus aureus.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 1..99  
FT /note= "X= any amino acid or stop codon"  
XX  
PN W09926475-A1.  
XX  
PD 03-JUN-1999.  
XX  
PF 23-NOV-1998; 98WO-US24857.  
XX  
PR 25-NOV-1997; 97US-0978456.  
XX  
PA (SMIK ) SMITHKLINE BEECHAM CORP.

```

XX Black MT, Burnham MKR, Fedon JC, Hodgson JE, Knowles DJC;
PI Lonetto MA, Nicholas RO, Palmer LM, Pratt JM, Reichard RW;
PI Rosenberg M, Trani CM, Ward JM, Warren RL;
XX
XX WPI: 1999-347572/29.
DR N-PSDB; X59910.
XX
XX New Staphylococcus aureus polypeptide and polynucleotide useful in
PT the treatment of gastric ulcer and gastritis
XX
XX Claim 23; Page 6; 48pp; English.
XX
XX The present sequence is encoded by the ORF of Staphylococcus aureus
CC pyrimidine deaminase and pyrimidine reductase (ribG) DNA. Staphylococcus
CC aureus ribG and its antagonists are used to treat individuals in
CC need of them. Disease related to expression or activity of ribG can
CC be determined by analysing the nucleic acid sequence encoding ribG
CC or detecting the ribG polypeptide in a sample. ribG can also be used
CC to identify antagonists or agonists. ribG, or its related nucleic acid,
CC also has use as a vaccine to induce an immunological response in an
CC animal. Antimicrobial compounds (e.g. agonists and antagonists of ribG),
CC especially broad-spectrum antibiotics, may be of use in the treatment
CC of Helicobacter pylori infection. This should decrease the advent of
CC H. pylori-induced cancers, such as gastrointestinal carcinoma. The
CC treatment should also cure gastric ulcers and gastritis.
XX
XX Sequence 99 AA;
SQ
Query Match 25.2%; Score 451; DB 20; Length 99;
Best Local Similarity 95.6%; Pred. No. 4.8e-35;
Matches 87; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 MDYAIOLANMVQGTGVNPPVGVAVVNEGRIVGIGAHLRKGDHAEVQALDMAQNAEGA 60
Db 1 mdyaiglpmmvgxtgvnppvgavvnegrivgigahlrkgdhaevqaldmaqnaega 60
QY 61 TIYITLPCSHFGSTPPCVNKKIIDCKIAKV 91
Db 61 tiyitlpcshfgstppcvnkkiiidckiaxv 91
RESULT 8
Y35519
ID Y35519 standard; Protein; 376 AA.
XX
AC Y35519;
XX
XX 13-SEP-1999 (first entry)
XX
XX Chlamydia pneumoniae transmembrane protein sequence.
XX
XX Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
KW sinusitis; purulent otitis media; erythema nodosum; pharyngitis;
KW vaccine; neutralising epitope.
XX
XX Chlamydia pneumoniae.
OS
XX
XX WO9927105-A2.
XX
XX 03-JUN-1999.
XX
XX 20-NOV-1998; 98WO-IB01890.
XX
XX 04-NOV-1998; 98US-0107078.
XX
XX 21-NOV-1997; 97FR-0014673.
XX
XX (CEST ) GENSET.
XX
XX Griffais R;
XX
XX WPI: 1999-357842/30.
XX

```

---

```

XX Genome sequence of Chlamydia pneumoniae
PT
XX Page 1273-1274; Disclosure; 1912pp; English.
XX
XX Y34584-Y35879 represent the proteins encoded by all the open reading
CC frames in the complete genome (see X91990) of Chlamydia pneumoniae.
CC C. pneumoniae causes respiratory disease such as pneumonia and
CC bronchitis and is thought to be a contributing factor in heart
CC disease, sarcoidosis, sinusitis, purulent otitis media, erythema
CC nodosum or pharyngitis. The polypeptides encoded by the open reading
CC frames of the C. pneumoniae genome (see Y34584-Y35879) can be used in
CC immunogenic compositions as vaccines. Vectors containing C. pneumoniae
CC nucleotide sequences can also be used as immunogenic compositions,
CC especially where the vector directs the expression of a neutralising
CC epitope of C. pneumoniae.
XX
XX Sequence 376 AA;
SQ
Query Match 24.3%; Score 435; DB 20; Length 376;
Best Local Similarity 31.2%; Pred. No. 1.1e-32;
Matches 112; Conservative 65; Mismatches 152; Indels 30; Gaps 9;
QY 1 MDYAIOLANMVQGTGVNPPVGVAVVNEGRIVGIGAHLRKGDHAEVQALDMAQNAEGA 60
Db 12 mrraiegekggritappnpwgcvvvqenriigefghayagpphaeeaiqnasmpisgs 71
QY 61 TIYITLPCSHFGSTPPCVNKKIIDCKIAKVYA--TKDNSLDTHGDETURAHGIEVEC-V 117
Db 72 dvuyvslepchsigscppcanllkhkvsrfvalvdppkavagqgiamlrqagivvygi 131
QY 118 DDERASQLYQDFEFKAKAKOLPQITVKVSASLDGKOANDNGSQWITNKVKODVYKLRHR 177
Db 132 gseasqaslpvlyqrthnfpwtlksaavdgqvadsgqskwtcpcarhadvkrlae 191
QY 178 HDAVLTGRTVELDDPQYTTTRIQDG----KNPIKVLTSKSGNIHFNQIYQDESTPIWY 233
Db 192 sqailvgstrvlssddpwtlcarpqgmlypkqlrvldsrsgvppskvfdktsptlyvt 251
QY 234 TENPNLTSTQTHIEIYIYKSCDLTFT-----LHNLV-----KRGVGTLLVEAGPTT 279
Db 252 ter----cpenyikv--ldslavpvltestpsgvdhkhvveylaqkklqlvveggttl 305
QY 280 -TSEFSIYVIDEFILYAPKLIIGSGNYQFYQTNVIEIPDANQPFIVHSELLNONVKL 337
Db 306 htsllkerfvnslvlysgpmiigdkrplvgvlgnlle--saspltkssqilgnslkv 362
RESULT 9
W55031
ID W55031 standard; Protein; 410 AA.
XX
AC W55031;
XX
XX 12-OCT-1998 (first entry)
XX
XX Actinobacillus pleuropneumoniae rib-specific deaminase.
XX
XX Vaccine; porcine pleuropneumonia; attenuation; riboflavin;
KW vitamin B2; ribGBAH operon; RibG protein; deaminase;
KW Pasteurella multocida; Pasteurella haemolytica;
KW Haemophilus parasuis; Haemophilus influenzae; Haemophilus ducreyi;
KW bronchopneumonia; fibrinous pleuritis.
XX
XX Actinobacillus pleuropneumoniae ISU178 (serotype 5).
OS
XX
XX WO9818917-A2.
XX
XX 07-MAY-1998.
XX
XX 28-OCT-1997; 97WO-US19649.
XX

```





```

DE Sinorhizobium meliloti RibD.
XX Plant growth; photosynthesis; alfalfa; riboflavin synthase;
KW riboflavin synthesis; RibD.
XX Sinorhizobium meliloti.
XX WO200029607-A1.
XX 25-MAY-2000.
XX 17-NOV-1999; 99WO-US27318.
XX 17-NOV-1998; 98US-0193600.
XX 17-NOV-1998; 98US-0193801.
XX 06-AUG-1999; 99US-0369955.
XX (REGC ) UNIV CALIFORNIA.
XX Phillips DA, Joseph CM, Sanborn JR, Yang G;
XX WPI: 2000-387815/33.
XX N-PSDB; A27366.
XX Enhancing plant growth by applying an agent comprising lumichrome or
XX lumichrome-releasing microorganism or growing plants in hydroponic
XX culture system or in a medium comprising lumichrome-releasing
XX microorganism .
XX Claim 61; Page 80-81; 86pp; English.
XX The present sequence is the Sinorhizobium meliloti RibD protein,
XX which is involved in riboflavin synthesis. Riboflavin is important
XX in plant growth and photosynthesis, and the protein and gene encoding it
XX can be used to create genetically engineered microorganisms which produce
XX an increased amount of the compound, which in turn increases plant
XX growth. In addition to being applied to the growing plant, the bacteria
XX can also be applied to the seed before planting. The bacteria can be
XX applied to any plant and any part of the plant.
XX Sequence 401 AA;
SQ
Query Match 19.9%; Score 356; DB 21; Length 401;
Best Local Similarity 27.4%; Pred. No. 3e-25;
Matches 101; Conservative 59; Mismatches 155; Indels 54; Gaps 9;
QY 1 MDYAIQLANWVGQTGVNPPVGVAVVNEGRIVGIGHLRKGDKHAEVQALDMAQQNAEGA 60
DB 12 maalrlarlnlgtstnpsvgcivnkgtlvgravtappgrphaetqalaesagekarga 71
QY 61 TIVITLPCSHFGSTPPCVNKIIDCKIAKVYA--TKDNLDPHGDETURAHCIEVEC-V 117
DB 72 tayvalcpshhktppcadalliasgvrvvvsildpdervagrgvmmrlrgaidvldigt 131
QY 118 DDERASQYODFFKAKAKQLPQITVKVSASLDGKQANDNGSQ--WITNKEVKODVYKLRH 176
DB 132 lheeggrvleaylmrqrkrphvtlklavsadg-migrgeqgvrqisgavstaqvqlra 190
QY 177 RHDVLTGRTVELDPOYTRTQ--DGNPKIVILSKSGNIHFNOQIYQD-ESTPIWIY 233
DB 191 etdailvgitadaddpeltvrmplgeerspvrivldrldlplesklvrtardvpllv 250
QY 234 TENPN-----LTSNPTHTETI----- 249
DB 251 tgadagypscheqagarradegssiplsagstrepvleaplaeaagtddarrgaleaaga 310
QY 250 -YLKSCDLTTLILNLKRGVGTLLVEAGPTTTFSEF-SIYVIDEILYAPKLLIGSGNYQ 307
DB 311 ellsaditpallaasrgissllvegaraarsfildadldvdriliftppaalgeggiss 370
QY 308 FYQTNVDVIE 316

```

---

```

Db 371 pfqrtsvpe 379
RESULT 12
W81550
ID W81550 standard; Protein; 249 AA.
XX W81550;
XX 01-MAR-1999 (first entry)
XX Bacillus subtilis rib-specific deaminase.
XX Riboflavin; vitamin B2; deaminase; rib operon.
XX Bacillus subtilis strain 168.
XX US5837528-A.
XX 17-NOV-1998.
XX 06-FEB-1995; 95US-0384626.
XX 11-SEP-1990; 90US-0581048.
XX 22-JUN-1989; 89US-0370378.
XX 21-APR-1992; 92US-0873572.
XX 06-FEB-1995; 95US-0384626.
XX (HOFF ) HOFFMANN LA ROCHE INC.
XX Perkins JB, Pero JG, Sloma A;
XX WPI: 1999-023456/02.
XX N-PSDB; V71600.
XX Nucleic acid encoding proteins involved in bacterial synthesis of
XX riboflavin - also mutant or recombinant bacteria producing
XX riboflavin at high levels
XX Example 7; Fig 3A-S; 62pp; English.
XX This is the amino acid sequence of a 44 kDa rib-specific deaminase
XX encoded by open reading frame 5 (ORF5) of the Bacillus subtilis rib
XX operon (see V71600). The deaminase is predicted to catalyze the
XX formation of 5-amino-6-(ribosylamino)-2,4(1H,3H)-pyrimidin-4(3H)-
XX 5'-phosphate from 2,5-diamino-6-(ribosylamino)-4(3H)-pyrimidin-4(3H)-
XX 5-phosphate, and shows 39% identity to the deoxycytidylate
XX deaminase of E. coli phage T2. The invention provides vectors and
XX recombinant bacteria for overproducing riboflavin, in which nucleic
XX acids overproducing riboflavin biosynthetic proteins (see
XX W81549-55) are introduced in the chromosome of the host organism at
XX multiple sites or in multiple copies. The rib operon, e.g.,
XX comprising ORFs 2-5 and the beta-riboflavin synthase gene, is used
XX to make such recombinant bacteria. The rib operon can be
XX deregulated by replacing its control regions with sequences that
XX allow constitutive or unregulated expression. The recombinant
XX bacteria produce at least 10 g/l riboflavin.
XX Sequence 249 AA;
SQ
Query Match 15.5%; Score 277; DB 20; Length 249;
Best Local Similarity 29.4%; Pred. No. 3.8e-18;
Matches 74; Conservative 52; Mismatches 100; Indels 26; Gaps 7;
QY 107 LRAHGTIEV-ECVDDERASQYQDFFKAKAKQLPQITVKVSASLDGKQANDNGSQWITNK 165
DB 2 mkeagievregiladgaerlnekfilhfmrtgipyvtlkaasldgklatstgdskwitse 61
QY 166 EVKQDVKYKLRHHRDAVLTGRTTVELDDPOYTRTQD--GKNPKIVILSKSGNIHFNOQIYQ 224
DB 62 aarqdaqykrkthqslivgvtgkdnpsltcrlpnvtkgprvrltdvlsipadakvic 121

```

Qy	166	EVKDDVYKLRHRHDAVLTGRTRVELDPPQVYTRIDQG---- <td>221</td>	221
		:        : :        :     :   : :   :	
Db	66	laradvgklraesqaiivgartvclndprlsarfphgdlyerqplrvvdsrgtvplesr	125
Qy	222	IYQDESTPIWYIENPNTSQTHIELIY-LK-----SCDLTTLIHLNLYKRGV	268
		:      :   :   :   :	
Db	126	vf-dlsqgstlfa-----ttgcpckeylqklklgvevwessshqvdikglrylaergc	179
Qy	269	GTLIVEAGPTTSEF-SIYYIDFELIYAPKLIGGSGNYQFYQTNVDVIEIPDANQFPIVH	327
		:                  :   :	
Db	180	lqvivegaqlhsafwqklnvagnviywgpkfkgdqgqpmldlq--lsivtaehwrite	237
Qy	328	SELLNQNVK 336	
		:   :	
Db	238	tslvrdsvk 246	
RESULT	14		
Y37428			
ID	Y37428	standard; Protein; 170 AA.	
XX	XX		
AC	Y37428;		
DT	07-OCT-1999	(first entry)	
XX	XX		
DE	XX	Amino acid sequence of a Chlamydia trachomatis protein.	
XX	XX		
KW	KW	Vaccine; eye disease; conventional trachoma; nonendemic trachoma;	
KW	KW	paratrachoma; inclusion conjunctivitis; genital disease; perihepatitis;	
KW	KW	nongonococcal urethritis; epididymitis; cervicitis; salpingitis;	
KW	KW	bartholinitis; pneumonia; venereal lymphogranulomatosis.	
XX	XX	Chlamydia trachomatis.	
OS	XX		
XX	XX	WO928475-A2.	
PN	XX		
XX	XX	10-JUN-1999.	
PD	XX		
PF	XX	27-NOV-1998; 98WO-IB01939.	
PR	XX	04-NOV-1998; 98US-0107077.	
PR	XX	28-NOV-1997; 97FR-0015041.	
PR	XX	17-DEC-1997; 97FR-0016034.	
XX	XX	(GEST ) GENSET.	
PA	XX		
XX	XX	Griffais R;	
PI	XX		
XX	XX	WPI; 1999-371125/31.	
XX	XX		
PT	XX	Genome sequence of Chlamydia trachomatis	
PS	XX	Disclosure; Page 1127-1128; 1755pp; English.	
XX	XX		
CC	XX	Y36754-Y37949 are encoded by open reading frames (ORFs) of the genome	
CC	XX	Chlamydia trachomatis (see Z01425). The polypeptides can be used as	
CC	XX	vaccines against Chlamydia trachomatis. Antisense and ribozyme sequen	
CC	XX	can also be used to control growth of the microorganism. Chlamydia	
CC	XX	trachomatis is responsible for a large number of diseases, e.g. eye	
CC	XX	diseases such as conventional trachoma, nonendemic trachoma,	
CC	XX	paratrachoma, and inclusion conjunctivitis; genital diseases such as	
CC	XX	nongonococcal urethritis, epididymitis, cervicitis, salpingitis,	
CC	XX	perihepatitis, bartholinitis; pneumopathy in breast feeding infants;	
CC	XX	and venereal lymphogranulomatosis. The polypeptides of the invention	
CC	XX	may be of use in treating these diseases.	
XX	XX		
XX	XX	Sequence 170 AA;	

Query Match 10.8%; Score 194; DB 20; Length 170;

Best Local Similarity 48.1%; Pred. No. 1.3e-10;

Matches 38; Conservative 11; Mismatches 30; Indels 0; Gaps









```

: APPLICANT: Doval, Jose Luis Revuelta
: APPLICANT: Serna, Maria Jose Buitago
: APPLICANT: Gareiss, Maria Angeles
: TITLE OF INVENTION: Riboflavin biosynthesis in fungi
: NUMBER OF SEQUENCES: 12
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Keil & Weinkauff
: STREET: 1101 Connecticut Avenue
: CITY: Washington
: STATE: D.C.
: COUNTRY: USA
: ZIP: 20036

```

REF: 20030  
COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage  
COMPUTER: IBM compatible  
OPERATING SYSTEM: MS-DOS version 7.0

SOFTWARE: WORDPERFECT VERSION 5.1  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08

FILING DATE: 24-SEP-

**CLASSIFICATION: 435**

CLASSIFICATION: C 12 N 15/53

CLASSIFICATION: C 12 N 15/54

CLASSIFICATION: C 12 N 15/35  
CLASSIFICATION: C 12 N 15/81

CLASSIFICATION: C 12 N 15/81  
CLASSIFICATION: C 12 N 1/19

CLASSIFICATION: C 12 N 1/15  
CLASSIFICATION: C 12 P 25/00

CLASSIFICATION: C 12 R 1:865

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/IN

FILING DATE: 15-MAR-1995

INFORMATION FOR SEQ ID NO:

SEQUENCE CHARACTERISTICS  
LENGTH. 609 amino acids

LENGTH: 003 amino acid  
TYPE: amino acid

TOPOLOGY: linear

; MOLECULE T

Query Match	6.6%	Score 117.5;	DB 2;	Length 609;
Best Local Similarity	30.5%;	Pred. No. 0.00039;		
Matches 39: Conservative	17: Mismatches 43:	Indels 29:	Gaps 8:	

OV 19 PP-----VGAVVNEGRIVGIG-AHLRKGDKHAEOALD--MAOONAE---GATIVYT 65

	:	:	:	:	:	:
—	—		: :	:	:	
—	—	—	—	—	—	—
—	—	—	—	—	—	—
—	—	—	—	—	—	—

66 LEPCS-HFGSTPPCVNKIIDCKIAKVYA-----TKDNS-----LDTHGDETLRAH 110

[illegible]

Qv 111 GIEVECD 118

0  
 4  
 4  
 1 ..  
 2 —  
 2 —  
 2 ..  
 2 —  
 2  
 2  
 2  
 2

RESULT 8  
US-08-133-248-8  
; Sequence 8, Application US/08133248  
; Patent No. 5525714  
; GENERAL INFORMATION:

```

; INFORMATION FOR SEQ ID NO: 8:
;
; SEQUENCE CHARACTERISTICS:
;
; LENGTH: 770 amino acids
;     TYPE: amino acid
;     TOPOLOGY: linear
;
; MOLECULE TYPE: protein
US-08-133-248-8

```

Query Match

OV 12 OGOTGVNPPVGAVV-----VNEGRTV-----GIGAHLRKGDKHAEVOALDMAOONAEFGA 60

Item	Unit	Quantity	Value	Remarks
1. Cement	kg	100	100	
2. Sand	kg	200	200	
3. Gravel	kg	300	300	
4. Water	kg	400	400	
5. Labor	hr	10	10	
6. Transport	km	5	5	
7. Fuel	liters	10	10	
8. Tools	set	1	1	
9. Materials	kg	50	50	
10. Other	kg	10	10	
<b>Total</b>			<b>1000</b>	

0y 61 TTVTTLERCSH-----FGSTPBCVNKIIDCKIAKVVYATKDNISIDTHGDETRAH-- 110

27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100

07 111 --GTFVECVDDERASOLYDDEFA--KAKOI BOTYKV--SASIDCKOBNDCO 159

QY III STEVE C DDERAS Q T I Q D F F R A K A R Q D E F Q I I V R V S A S L D G R Q A N D N G Q 138

Q#	150	SOWITNKFKVKNQVNVKI	PIRURPAVIT	TCRRPWIVET	DDRRQVETPIQ	-----	300
Q#	150	SOWITNKFKVKNQVNVKI	PIRURPAVIT	TCRRPWIVET	DDRRQVETPIQ	-----	300

QY 139 SQWIIINKEVRQDVIRLKHKHDAVLIGRKRIVELDDFQITIRIQ----- 200

07 201 ---DCVNDTVWTT CVCCNTHUENAO TVOCESTDTVTVEENDNT EGNCHTET T VY WCOBY CES

QY ZOI ---DGRNPKVLLSKSGNIHFNQIYQDESTPIWIYITENPNLTNSQTHIELIYLRKSDL 256

0-7 957 UNIT 762

00 757 T.I.L.H.N.L 263  
.:|.:.:

```

RESULT      9
US-08-231-940-1
; Sequence 1, Application US/08231940
; Patent No. 5550216
; GENERAL INFORMATION:
; APPLICANT: MIYAZAKI, Kaoru
; TITLE OF INVENTION: GELATINASE A INHIBITOR
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEWMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA

```

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/231,940  
FILING DATE: 25-APR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 5-120457  
FILING DATE: 26-APR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 6-62129  
FILING DATE: 08-MAR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: NEIMARK, Sheridan  
REGISTRATION NUMBER: 20,520  
REFERENCE/DOCKET NUMBER: M17AZAKI-4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-628-5197



```
;
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 770 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-231-940-1

Query Match          5.2%; Score 93; DB 1; Length 770;
Best Local Similarity 21.5%; Pred. No. 0.28;
Matches 66; Conservative 41; Mismatches 96; Indels 104; Gaps 14;

QY 12 QGQTGVNPPVGAVV-----VNEGRIV-----GIGAHLRKDGKHAEQALDMAQNAEGA 60
   | : | | | | : | | : | | | | : | | | | : | | | | : | | | | :
Db 294 QAETG---PCRAMISRWFVDTGKCAPFFYGGCGGNRRNFD--TEECMAVCGSAMSQS 348
   : | | | | : | | | | : | | | | : | | | | : | | | | : | | | | :
Db 349 LLKTTQEPPLARDPVKLPPTTAASTDAVDKY-----LTPGCDENEHAHFQ 392
   : | | | | : | | | | : | | | | : | | | | : | | | | : | | | | :
QY 111 --GIEVECVDDERASQYQDFEKA--KAKQLPQITVKV-----SASLDGKQANDNGQ 158
   : | | | | : | | | | : | | | | : | | | | : | | | | : | | | | :
Db 393 KAKERLEAKHRERMSQVMREWEAEERQAKNLPKADKAVIQHFQEKVESLEQEAAN---- 448
   : | | | | : | | | | : | | | | : | | | | : | | | | : | | | | :
QY 159 SOWITNKEVKQDVYKLRHRHDAVLTRGTVELDDPQYVTRIQ----- 200
   | : | | | | : | | | | : | | | | : | | | | : | | | | : | | | | :
Db 449 -----ERQQLVETHMARVEMLNDRRRLALEN--YITALQAVPPRPHVFNMLKKYVR 499
   : | | | | : | | | | : | | | | : | | | | : | | | | : | | | | :
QY 201 ----DGKNPIKVLKSGNIHFNQIYQDESTPIWIYTNPNLTSNTHIETIYILKSCDL 256
   | : | | | | : | | | | : | | | | : | | | | : | | | | : | | | | :
Db 500 AEQDRQHTLK-----HFEHVRMVDPKKAAQIRSQ-----VMTHLRVIVERNQS 544
   : | | | | : | | | | : | | | | : | | | | : | | | | : | | | | :
QY 257 TTILNL 263
Db 545 LSLLYNV 551

RESULT 10
US-08-641-774-1
; Sequence 1, Application US/08641774
; Patent No. 5843695
; GENERAL INFORMATION:
; APPLICANT: MIYAZAKI, Kaoru
; TITLE OF INVENTION: BETA-AP DECOMPOSING AGENT
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/641,774
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/232,474
; FILING DATE: 25-APR-1994
; APPLICATION NUMBER: JP 5-122207
; FILING DATE: 26-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-51133
; FILING DATE: 08-MAR-1994
; ATTORNEY/AGENT INFORMATION:
```

```
;
; NAME: NEIMARK, Sheridan
; REGISTRATION NUMBER: 20,520
; REFERENCE/DOCKET NUMBER: MIYAZAKI=3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 770 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-641-774-1

Query Match          5.2%; Score 93; DB 2; Length 770;
Best Local Similarity 21.5%; Pred. No. 0.28;
Matches 66; Conservative 41; Mismatches 96; Indels 104; Gaps 14;

QY 12 QGQTGVNPPVGAVV-----VNEGRIV-----GIGAHLRKDGKHAEQALDMAQNAEGA 60
   | : | | | | : | | : | | | | : | | | | : | | | | : | | | | :
Db 294 QAETG---PCRAMISRWFVDTGKCAPFFYGGCGGNRRNFD--TEECMAVCGSAMSQS 348
   : | | | | : | | | | : | | | | : | | | | : | | | | : | | | | :
QY 61 TIYITLPCSH-----FGSTPPCVNKKIIDCKIAKVYVYATKDNSLDTHGDETLRAH-- 110
   : | | | | : | | | | : | | | | : | | | | : | | | | : | | | | :
Db 349 LLKTTQEPPLARDPVKLPPTTAASTDAVDKY-----LTPGCDENEHAHFQ 392
   : | | | | : | | | | : | | | | : | | | | : | | | | : | | | | :
QY 111 --GIEVECVDDERASQYQDFEKA--KAKQLPQITVKV-----SASLDGKQANDNGQ 158
   : | | | | : | | | | : | | | | : | | | | : | | | | : | | | | :
Db 393 KAKERLEAKHRERMSQVMREWEAEERQAKNLPKADKAVIQHFQEKVESLEQEAAN---- 448
   : | | | | : | | | | : | | | | : | | | | : | | | | : | | | | :
QY 159 SOWITNKEVKQDVYKLRHRHDAVLTRGTVELDDPQYVTRIQ----- 200
   | : | | | | : | | | | : | | | | : | | | | : | | | | : | | | | :
Db 449 -----ERQQLVETHMARVEMLNDRRRLALEN--YITALQAVPPRPHVFNMLKKYVR 499
   : | | | | : | | | | : | | | | : | | | | : | | | | : | | | | :
QY 201 ----DGKNPIKVLKSGNIHFNQIYQDESTPIWIYTNPNLTSNTHIETIYILKSCDL 256
   | : | | | | : | | | | : | | | | : | | | | : | | | | : | | | | :
Db 500 AEQDRQHTLK-----HFEHVRMVDPKKAAQIRSQ-----VMTHLRVIVERNQS 544
   : | | | | : | | | | : | | | | : | | | | : | | | | : | | | | :
QY 257 TTILNL 263
Db 545 LSLLYNV 551

RESULT 11
US-08-104-165-3
; Sequence 3, Application US/08104165
; Patent No. 5877015
; GENERAL INFORMATION:
; APPLICANT: HARDY, John Anthony
; APPLICANT: GOATE, Allison Mary
; APPLICANT: MULLAN, Michael John
; APPLICANT: CHARTIER-HARLIN, Marie-Christine
; APPLICANT: OWEN, Michael John
; TITLE OF INVENTION: Test and Model for Alzheimer's Disease
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: 379 Lytton Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: US
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/104,165
; FILING DATE: 21-JAN-1992
; CLASSIFICATION: 435
```

```
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 9101307.8
; FILING DATE: 21-JAN-1991
; APPLICATION NUMBER: 9118445.7
; FILING DATE: 28-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Liebeschuetz, Joe
; REGISTRATION NUMBER: 37,505
; REFERENCE/DOCKET NUMBER: 16163-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 770 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
US-08-104-165-3

Query Match          5.2%; Score 93; DB 2; Length 770;
Best Local Similarity 21.5%; Pred. No. 0.28;
Matches 66; Conservative 41; Mismatches 96; Indels 104; Gaps 14;

QY 12 QGOTGVNPPVGAIV-----VNEGRIV-----GIGAHLRKRGDKHAEVQALDMAQQNAEGA 60
   | : | | | : | | | : | | | : | | | : | | | : | | | : | | | : | | | :
Db 294 QAETG---PCRAMISRWFYDTEGKCAPFYGCGGNRRNFD---TEEYCMVCGSAMSOS 348

QY 61 TYITILEPCSH-----FGSTPPCVNKIIDCKIAKVYVATKDNSLDTHGDETTLRAH-- 110
   | : | | | : | | | : | | | : | | | : | | | : | | | : | | | : | | | :
Db 349 LKTTQOEPLARDPVKLTPTTAASTPDAVDKY-----LETPGDENEHAHFQ 392

QY 111 --GIEVECVDDERASQYQDFKA--KAKOLPOITVKV-----SASLDGKOANDNGQ 158
   | : | | | : | | | : | | | : | | | : | | | : | | | : | | | : | | | :
Db 393 KAKERLEAKHRERMSQVMREWEAEARQAKNLPKADKKAVIQHFQEKVESLEQEAAN---- 448

QY 159 SQWITNKEVKQDVYKLRHRHDVLTGRRTVELDDPOYTTRIQ----- 200
   | : | | | : | | | : | | | : | | | : | | | : | | | : | | | : | | | :
Db 449 -----ERQQLVETHMARVEAMLNDRRLALEN--YITALQAVPPRPHVFNMLKKYYR 499

QY 201 -----DGKNPIKVLKSGNIHFNQIYQDESTPIWYITENPNLTSNQTIEIYYLKSDDL 256
   | : | | | : | | | : | | | : | | | : | | | : | | | : | | | : | | | :
Db 500 AEQKDRQHTLK-----HFEHVRMVDPKKAAQIRSQ-----VMTHLRVYIERMNQS 544

QY 257 TTILHNL 263
   | : | | | : | | | : | | | : | | | : | | | : | | | : | | | : | | | :
Db 545 LSLLYNV 551

RESULT 12
US-08-464-250-3
; Sequence 3, Application US/08464250
; Patent No. 6107542
; GENERAL INFORMATION:
; APPLICANT: HARDY, John Anthony
; APPLICANT: GOATE, Alison Mary
; APPLICANT: MULLAN, Michael John
; APPLICANT: CHARTIER-HARLIN, Marie-Christine
; APPLICANT: OWEN, Michael John
; TITLE OF INVENTION: Test and Model for Alzheimer's Disease
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: 379 Lytton Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: US
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC compatible
```

```
;
; OPERATING SYSTEM: PC-DOS/MS-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/464,250
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/104,165
; FILING DATE: 21-JAN-1992
; APPLICATION NUMBER: 9101307.8
; FILING DATE: 21-JAN-1991
; APPLICATION NUMBER: 9118445.7
; FILING DATE: 28-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Liebeschuetz, Joe
; REGISTRATION NUMBER: 37,505
; REFERENCE/DOCKET NUMBER: 16163-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 770 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
US-08-464-250-3

Query Match          5.2%; Score 93; DB 3; Length 770;
Best Local Similarity 21.5%; Pred. No. 0.28;
Matches 66; Conservative 41; Mismatches 96; Indels 104; Gaps 14;

QY 12 QGOTGVNPPVGAIV-----VNEGRIV-----GIGAHLRKRGDKHAEVQALDMAQQNAEGA 60
   | : | | | : | | | : | | | : | | | : | | | : | | | : | | | : | | | :
Db 294 QAETG---PCRAMISRWFYDTEGKCAPFYGCGGNRRNFD---TEEYCMVCGSAMSOS 348

QY 61 TYITILEPCSH-----FGSTPPCVNKIIDCKIAKVYVATKDNSLDTHGDETTLRAH-- 110
   | : | | | : | | | : | | | : | | | : | | | : | | | : | | | : | | | :
Db 349 LKTTQOEPLARDPVKLTPTTAASTPDAVDKY-----LETPGDENEHAHFQ 392

QY 111 --GIEVECVDDERASQYQDFKA--KAKOLPOITVKV-----SASLDGKOANDNGQ 158
   | : | | | : | | | : | | | : | | | : | | | : | | | : | | | : | | | :
Db 393 KAKERLEAKHRERMSQVMREWEAEARQAKNLPKADKKAVIQHFQEKVESLEQEAAN---- 448

QY 159 SQWITNKEVKQDVYKLRHRHDVLTGRRTVELDDPOYTTRIQ----- 200
   | : | | | : | | | : | | | : | | | : | | | : | | | : | | | : | | | :
Db 449 -----ERQQLVETHMARVEAMLNDRRLALEN--YITALQAVPPRPHVFNMLKKYYR 499

QY 201 -----DGKNPIKVLKSGNIHFNQIYQDESTPIWYITENPNLTSNQTIEIYYLKSDDL 256
   | : | | | : | | | : | | | : | | | : | | | : | | | : | | | : | | | :
Db 500 AEQKDRQHTLK-----HFEHVRMVDPKKAAQIRSQ-----VMTHLRVYIERMNQS 544

QY 257 TTILHNL 263
   | : | | | : | | | : | | | : | | | : | | | : | | | : | | | : | | | :
Db 545 LSLLYNV 551

RESULT 13
US-08-123-702-4
; Sequence 4, Application US/08123702
; Patent No. 5604131
; GENERAL INFORMATION:
; APPLICANT: Wadsworth, Samuel
; APPLICANT: Snyder, Benjamin
; APPLICANT: Reddy, Vermuri, B.
; APPLICANT: Wei, Chamer
; TITLE OF INVENTION: A cDNA Genomic Hybrid Sequence Encoding APP770
; Patent No. 5604131
; TITLE OF INVENTION: Containing a Genomic DNA Insert of the KI and OX-2 Regions
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
```

COUNTRY: US  
ZIP: 94301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy Disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/104,165  
FILING DATE: 21-JAN-1992  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 9101307.8  
FILING DATE: 21-JAN-1991  
APPLICATION NUMBER: 9118445.7  
FILING DATE: 28-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Liebeschuetz, Joe  
REGISTRATION NUMBER: 37,505  
REFERENCE/DOCKET NUMBER: 16163-000100  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 326-2400  
TELEFAX: (415) 326-2422  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 751 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-104-165-2

Query Match 5.1%; Score 91.5; DB 2; Length 751;  
Best Local Similarity 22.1%; Pred. No. 0.39;  
Matches 66; Conservative 38; Mismatches 88; Indels 107; Gaps 14;

Qy	12	QGOTGVNPDVGAVV	-----VNEGRIV-----	-GIGAHLRKGDKHAEVQALDMAQQA	50
Db	294	QAETG	--PCRAMISRWYEDVTGECAPFYGGCGGNRNFT	EETEEYCM	343
Qy	61	TYITILEPCSHFGSTPPCNKIIDCKIAKVVYATK	DNLSLDTHGDET	LAH-----GIEV	116
Db	344	AIPTTA	-----ASTPDAVDKY-----	LET	381
Qy	117	VDDERASQLYQDFFKA	-KAKOLPQITVKV-----	SASLDGQ	166
Db	382	KHREMSQYMRWEAEARQAKNPKADKAVIQHFQ	KEVSELEQAAN-----	E	430
Qy	167	VHQDVKVLRRHRDVAVLTRGTVELDDPOVTT	RIQ-----	DGN	204
Db	431	RQOLVETHMARVEMLNDRRRLALEN	-YITALQAVPPRPHRVFNWMLKYYVRAE	QKDROH	488
Qy	205	PKIVILKSGNIHFNOQIYQDSESTPIWITENPN	LTSNQTHTIEIYLKSCD	LTILHNL	263
Db	489	TLK	-----HFEBHRMVDPKKAAQIRSQ-----	VMT	532

RESULT 15  
US-08-422-333-2  
; Sequence 2, Application US/08422333  
; Patent No. 5912410  
; GENERAL INFORMATION:  
; APPLICANT: CORDELL, Barbara L.  
; TITLE OF INVENTION: TRANSGENIC NON-HUMAN MAMMAL DISPLAYING  
; TITLE OF INVENTION: THE AMYLOID-FORMING PATHOLOGY OF ALZHEIMER'S DISEASE  
; NUMBER OF SEQUENCES: 30  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Scios, Inc.  
; STREET: 2450 Bayshore Parkway  
; CITY: Mountain View  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94043

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/422,333
; FILING DATE: 13-APR-1995
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Shearer, Peter R.
; REGISTRATION NUMBER: 28,117
; REFERENCE/DOCKET NUMBER: 21900-28048.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 966-1550
; TELEFAX: (415) 968-2438
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 751 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-422-333-2

Query Match      5.1%; Score 91.5; DB 2; Length 751;
Best Local Similarity 22.1%; Pred. No. 0.39;
Matches 66; Conservative 38; Mismatches 88; Indels 107; Gaps 14;

QY 12 QGOTGVNPPVGAVV-----VNEGRIV-----GIGAHLRKGDKHAEVQALDMAQQNAEGA 60
Db 294 QAETG---PCRAMISRWFYDVTGKCAPFFYGGCGGNRNNDTEECMAVC-----GS 343
QY 61 TIYTILEPCSHFGSTPCVYNKIIDCKIAKVYATKDNSLDTHGDETLRAH---GIEVEC 116
Db 344 AIPPTA-----ASTPDVDKY-----LETPGDENEHAHFQKAKERLEA 381
QY 117 VDDERASQLYODFFKA--KAKQLPQITVKV-----SASLDGKQANDNGOSQWITNKE 166
Db 382 KHERMSQVMREWEAEERQAKNLPKADKRAVIOHFQEKVESLEQEAAN-----E 430
QY 167 VKQDVKLRHRHDAVLTGRTVELDDPQYTRIQ-----DGNK 204
Db 431 ROQLVETHMARVEAMLNDRRLALEN--YITALQAVPPRPHVFNMLKKYVRAEQKDROH 488
QY 205 PIKVILSKSGNIHFNQIYODESTPIWIYTNPNLTSTNTHIEIYLYKSCDLTTILHNL 263
Db 489 TLK-----HFEHVRMVDPKAAQIRSQ-----VMTHLRVIYERMNQSLSLDYNV 532
```

Search completed: November 14, 2000, 10:07:45  
Job time: 63 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 14, 2000, 10:06:42 ; Search time 24.32 Seconds  
(without alignments)  
892.409 Million cell updates/sec

Title: US-08-978-456-2  
Perfect score: 1789  
Sequence: 1 MDYAIQLANMVGQGVNPP.....FEIVHSELLNONVKLTLRKK 342  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 182106 seqs, 63460219 residues  
Total number of hits satisfying chosen parameters: 182106

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_65:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	547	30.6	361	PN0100	riboflavin-specific
2	545	30.5	356	C70313	riboflavin-specific
3	530	29.6	348	G72207	riboflavin-specific
4	495.5	27.7	372	H64103	riboflavin-specific
5	484	27.1	369	H81039	riboflavin-specific
6	480	26.8	369	H81984	probable diaminohy
7	470	26.3	367	S26201	riboflavin-specific
8	435	24.3	376	G72026	riboflavin-specific
9	434.5	24.3	339	F70901	probable riboflavi
10	422	23.6	357	G75552	riboflavin-specific
11	413.5	23.1	368	S74372	riboflavin-specific
12	411.5	23.0	396	D81472	riboflavin biosynt
13	408.5	22.8	375	E71479	riboflavin-specific
14	320	17.9	599	T12994	probable riboflavi
15	300	16.8	363	T10638	riboflavin biosynt
16	227.5	12.7	336	D81258	hypothetical prote
17	195	10.9	219	F69500	probable riboflavi
18	191	10.7	224	G64383	riboflavin-specific
19	184	10.3	216	C69129	riboflavin-specific
20	180	10.1	224	G70339	riboflavin-specific
21	163	9.1	213	S74011	probable riboflavi
22	146.5	8.2	178	F65033	hypothetical 20.0
23	142.5	8.0	151	G70377	conserved hypothet
24	142	7.9	148	D81879	probable cytosine
25	142	7.9	173	C64161	hypothetical prote
26	140	7.8	344	A64708	riboflavin biosynt
27	140	7.8	344	G71813	probable riboflavi
28	139	7.8	164	S74803	hypothetical prote
29	139	7.8	239	H81141	cytidine and deoxy

conserved hypothet  
probable deoxycyti  
hypothetical prote  
dCMP deaminase (EC  
dCMP deaminase (EC  
riboflavin biosynt  
hypothetical prote  
hypothetical prote  
probable cell cycl  
hypothetical prote  
cytosine deaminase  
probable cytosine  
dCMP deaminase hom  
RIB2 protein - yea  
hypothetical prote  
probable RNA polym

ALIGNMENTS

RESULT 1  
PN0100  
riboflavin-specific deaminase ribG - Bacillus subtilis  
C:Species: Bacillus subtilis  
C:Date: 04-Dec-1992 #sequence\_revision 06-Jan-1995 #text\_change 20-Jun-2000  
C:Accession: S45543; PN0100; E69692  
R:Sorokin, A.; Zumstein, E.; Azevedo, V.; Ehrlich, S.D.; Serror, P.  
submitted to the EMBL Data Library, November 1993  
A:Reference number: S45533  
A:Accession: S45543  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-361 <SOR>  
A:CROSS-references: EMBL:L09228; NID:g410114; PIDN:AAA67481.1; PID:g410125  
R:Mironov, V.N.; Perumov, D.A.; Krayev, A.S.; Stepanov, A.I.; Skryabin, K.G.  
Mol. Biol. (Mosk.) 24, 256-261, 1990  
A:Title: Unusual structure of Bacillus subtilis rib-operon regulatory region.  
A:Reference number: PN0100; MUID:90271920  
A:Accession: PN0100  
A:Molecule type: DNA  
A:Residues: 1-12 <MIR>  
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber  
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;  
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari,  
Nature 390, 249-256, 1997  
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gal  
lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M  
Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardino  
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau  
Y, M.; Ogawa, K.; Ogilwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portete  
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanl  
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Se  
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiya  
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida  
A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.  
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtili  
A:Reference number: A69580; MUID:98044033  
A:Accession: E69692  
A>Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-361 <KUN>  
A:CROSS-references: GB:299116; GB:AL009126; NID:g2634723; PIDN:CAB14260.1; PID:g26347  
A:Experimental source: strain 168  
C:Genetics:  
A:Gene: ribG  
C:Superfamily: Chlamydomphila pneumoniae riboflavin-specific deaminase

Query Match 30.6%; Score 547; DB 2; Length 361;  
Best Local Similarity 36.1%; Pred. No. 3.8e-34;  
Matches 130; Conservative 68; Mismatches 134; Indels 28; Gaps 8;





```

RESULT 7
S26201
riboflavin-specific deaminase (EC 3.5.4.-) - Escherichia coli
C:Species: Escherichia coli
C:Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 20-Apr-2000
C:Accession: S26201; F64770; S19178
R:Taura, T.; Ueguchi, C.; Shiba, K.; Ito, K.
Mol. Gen. Genet. 234, 429-432, 1992
A:Title: Insertional disruption of the nusB (ssyB) gene leads to cold-sensitive growth of Escherichia coli K-12.
A:Reference number: S26200; MUID:93024316
A:Accession: S26201
A:Molecule type: DNA
A:Residues: 1-367 <TAU>
A:Cross-references: EMBL:X64395; NID:g42147; PIDN:CAAA45735.1; PID:g581147
A:Experimental source: strain K-12
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Cohen, D.; Rose, D.J.; Mau, B.; Shaoh, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617
A:Accession: F64770
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-367 <BLAT>
A:Cross-references: GB:AE000148; GB:U00096; NID:gl786614; PIDN:AAC73517.1; PID:gl786616;
A:Experimental source: strain K-12, substrain MG1655
C:Genetics:
A:Gene: ribD; ribG
A:Start codon: GTG
C:Function:
A:Description: converts 2,5-diamino-6-(riboseylamino)-4(3H)-pyrimidinone 5'-phosphate into 2,5-diamino-6-aminouracil
A:Pathway: riboflavin biosynthesis
C:Superfamily: Chlamydomophila pneumoniae riboflavin-specific deaminase
C:Keywords: hydrolase; riboflavin biosynthesis; zinc
F:50,75,84/Binding site: zinc (His, Cys, Cys) #status predicted

Query Match 26.3%; Score 470; DB 2; Length 367;
Best Local Similarity 33.1%; Pred. No. 2.7e-28;
Matches 120; Conservative 58; Mismatches 135; Indels 50; Gaps 8;

QY 1 MDYAIQLANNVQQTGTPNPPVGVVYVNEGRIVGIGAHLRKDKHAEVQALDMAQQNAEGA 60
Db 7 MARALKLAQRGRFTTHPNPVGCVIVKDGIEVGEYHQRAGEPHAEVHALRMAGERAKGA 66

QY 61 TIVTILEPCSHFGSTPPCVNKKIIDCKIAKVVA--TKDNSLDTHGDETIRAHGIEVEC-V 112
Db 67 TAVTILEPCSHHGRTPCCDALLAAGVARVVASMQDPNPFQVAGRGYLRLOQAGIDVSHGL 126

QY 113 EVECVDDERASQLYQDFFKAKAKQLPQITVKVSASLDGKQANDNGSQWITNKKEVRQDQV 172
Db 127 MM-----SEAEQLNKGFLKMRRTGFPYQLKLGASLDGRMTAMASGESQWITSPQARRDQ 181

QY 173 KLRHHDVAUTGRTVELDDPQVTR-----IQQGKNPIKVLKSGNHN 219
Db 182 LLRAQSHAILTSATLADDPALTVRWSELDEQTAQLYPOQNLQRPIRIVIDSQNRVTPV 241

QY 220 QQIYQDESTPIWIYTNPN-----LTSNPHIEIILKSCDLTILHNLKRGV 268
Db 242 HRIVQPGFTWARTQDSREWPETVRTLPIPEKHGL-----DLVVMMLQGLKQOI 293

QY 269 GTLLVEAGPTTTSF-PSIYVIDEILYAPKLGSGNRYQFYQTNVDVIEIPDANQF---E 324
Db 294 NSIWEAGPTLAGALLQAGLVDELIVYIAPKLGSDAR-GLCTLPGLKLAADAPQFKFE 352

QY 325 IVH 327
Db 353 IRH 355

RESULT 8
G72026
riboflavin-specific deaminase CP0998 [imported] - Chlamydomophila pneumoniae (strains C
C:Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000
C:Accession: G72026; H81514
R:Kalan, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood,
Nature Genet. 21, 385-389, 1999
A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
A:Reference number: A72000; MUID:99206606
A:Accession: G72026
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-376 <ARN>
A:Cross-references: GB:AE001667; GB:AE001363; NID:g4377171; PIDN:AAD19009.1; PID:g437
A:Experimental source: strain CWL029
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heideberg, J.F.; White, O.; Hicke
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzbe
Nucleic Acids Res. 28, 1397-1406, 2000
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39
A:Reference number: A81500; MUID:20150255
A:Accession: H81514
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-376 <REA>
A:Cross-references: GB:AE002257; GB:AE002161; NID:g7189902; PIDN:AAF38776.1; PID:g718
A:Experimental source: strain AR39, HL cells
C:Genetics:
A:Gene: ribD; CP0998
C:Superfamily: Chlamydomophila pneumoniae riboflavin-specific deaminase

Query Match 24.3%; Score 435; DB 2; Length 376;
Best Local Similarity 31.2%; Pred. No. 1.3e-25;
Matches 112; Conservative 65; Mismatches 152; Indels 30; Gaps 9;

QY 1 MDYAIQLANNVQQTGTPNPPVGVVYVNEGRIVGIGAHLRKDKHAEVQALDMAQQNAEGA 60
Db 12 MRRRAIEGKGRITAPPNPWGVGVVQVQENRIIGEFGHAYAGPHAEELAIQNSMPISGS 71

QY 61 TIVTILEPCSHFGSTPPCVNKKIIDCKIAKVVA--TKDNSLDTHGDETIRAHGIEVEC-V 117
Db 72 DVYVSLGPCSHFGSCPCANLLIKHKVSRVFVALVDPDPKVAQGLAMLRQAGIQVYVGI 131

QY 118 DDERASQLYQDFFKAKAKQLPQITVKVSASLDGKQANDNGSQWITNKKEVDYKLRHR 177
Db 132 GESEAQASLQPYLYQRTNFPMTILKSAASVDQGVADSGQSKSWITCPEARHVDGKLRAE 191

QY 178 HDVAULTGRTVELDDPQVTRIQDG----KNPIKVLKSGNHNHFNQIYQDESTPIWIY 233
Db 192 SQAILVGSRTVLSDDDPMLTARQPGMLYKQPLRVVLDSSRGVSPPTSKVFDKTSPTLYVT 251

QY 234 TENPNTSNOTHIEIILKSCDLTTI-----LHNLV----KRGVGTLLVEAGPTT 279
Db 252 TER---CPENIKV--LDSLDVPLLTESTPGVDLHKVYELAQKILQVIEGVTTL 305

QY 280 -TSEFSIYVIDEILYAPKLGSGNRYQFYQTNVDVIEIPDANQFIVHSELNLQNVL 337
Db 306 HTSLKLERFVNSLVLYSGPMILGDKRPLVGLGNLE--SASPLTLKSSQILGNSLKV 362

RESULT 9
F70901
probable riboflavin-specific deaminase - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C:Accession: F70901
R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon
Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd,
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno
A:Reference number: A70500; MUID:98295987
A:Accession: F70901
```



[illegible]

RESULT 12	
D81742	
riboflavin-specific deaminase TC0103 [imported] - Chlamydia muridarum (strain Nigg)	
C:Species: Chlamydia muridarum, Chlamydia trachomatis MoPn	
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 11-May-2000	
C:Accession: D81742	
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg, Nucleic Acids Res. 28, 1397-1406, 2000	
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.	
A:Reference number: A81500; MUID:20150255	
A:Accession: D81742	
A:Status: preliminary	
A:Molecule type: DNA	
A:Residues: 1-396 <FET>	
A:Cross-references: GB:AE002277; GB:AE002160; NID:g7190127; PIDN:AAF38983.1; PID:g719013	
A:Experimental source: strain Nigg (MoPn)	
C:Genetics:	
A:Gene: TC0103	
Query Match 23.0%; Score 411.5; DB 2; Length 396;	
Best Local Similarity 30.9%; Pred. No. 8.3e-24;	
Matches 104; Conservative 67; Mismatches 143; Indels 23; Gaps 8;	
QY 18	NPVGVAVVNEGRIVGIGAHLRKGDHAEVQALDMAQNAEGATIIYITLPCSHFGSTPP 77
DB 54	NPWVCVIVKNGCIIGRWKHGIGSPHAEVCAFDQDTSSLVAGADVIVVTLPCCHFGRTTP 113
QY 78	CVNKKIIDCKIAKVYA--TKDNLSDTHGDETLRAHGIIEVC-VDDERASQLYQDFPKA 134
DB 114	CVDLLIKSVSSVYIALLDPPRVCKRGVARLKEAGISVYVVGIGHEAKASLSQPLYLHQRE 173
QY 135	KOLPQITVVKVASLDGKQANDNCSOWITNKEVKQDVYKLRHRHDAVLGRRRTVELDDPQ 194
DB 174	TGLPWWVMKTAASLDGQTSRRGISQWISGEQARLDVGRLAESQAVIVGSRIVCLDNPR 233
QY 195	YTTRIODG----KNPIKVILSKSGNHFNQOIYODESTPIWIYTNPNLTSTNQTHIE 247
DB 234	LSARMPGSDGLYEROPLRVVVDSRGVPLDARVNPDSGNVLLATTE---QCSKEHIQKLE 290
QY 248	---IYLKS----CDLTTILHNLKRGVGTLLVEAGPTTISEF-SIYYIDEFILYAPKL 299
DB 291	DRGVEVWKSPPQDVLDKRLQLYLAERGCQLVLEGGARLHSAFWRHLLVNGAGVIYNGPKF 350
QY 300	IGSGSNVQFYQTNVDIEIPDANQFEIVHSELNQNPK 336
DB 351	LGDQSGPMLRLDQ--LCLDNAEHVKTKTFLVGDGSK 385
RESULT 13	
E71479	
probable riboflavin deaminase - Chlamydia trachomatis (serotype D, strain UW3/Cx)	
C:Species: Chlamydia trachomatis	
C:Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 20-Apr-2000	
C:Accession: E71479	
R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell, Science 282, 754-759, 1998	
A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trachomatis	
A:Reference number: A71570; MUID:99000809	
A:Accession: E71479	
A:Status: preliminary	
A:Molecule type: DNA	
A:Residues: 1-375 <ARN>	
A:Cross-references: GB:AE001343; GB:AE001273; NID:g3329176; PIDN:AAC68325.1; PID:g332918	
A:Experimental source: serotype D, strain UW-3/Cx	
C:Genetics:	
A:Gene: ribD	
C:Superfamily: Chlamydia pneumoniae riboflavin-specific deaminase	
Query Match 22.8%; Score 408.5; DB 2; Length 375;	
Best Local Similarity 31.8%; Pred. No. 1.3e-23;	

Matches 108; Conservative 58; Mismatches 145; Indels 29; Gaps 9;	
QY 18	NPVGVAVVNEGRIVGIGAHLRKGDHAEVQALDMAQNAEGATIIYITLPCSHFGSTPP 77
DB 29	NPWVCVIVKNGCVIGEGWHQIGSPHAEVCAVDQDKCSLEGAEEVFTLEPCCHFGRTTP 88
QY 78	CVNKKIIDCKIAKVYA--TKDNLSDTHGDETLRAHGIIEVC-VDDERASQLYQDFPKA 134
DB 89	CVDLLIKSVAAVYVGLDPPRVCKRGVARLQAAGIPVYVGVGSQEAFTSLQPLYLQRE 148
QY 135	KOLPQITVVKVASLDGKQANDNCSOWITNKEVKQDVYKLRHRHDAVLGRRRTVELDDPQ 194
DB 149	RGLPWWVMKTAASLDGQTSRRGISQWISGEQARLDVGRLAESQAVIVGARTVCLDNPR 208
QY 195	YTTRIODG----KNPIKVILSKSGNHFNQOIYODESTPIWIYTNPNLTSTNQTHIEIY 250
DB 209	LSARPPHGDGLYEROPLRVVVDSRGVPLSRVVF-DLSSGSTLFA----TTQOCPEYIQ 262
QY 251	-LK-----SCDLTTILHNLKRGVGTLLVEAGPTTISEF-SIYYIDEFILY 296
DB 263	KLKDLGVEVWESSSHQVDLKLRLYLAERGCQLVLEGGARLHSAFWRHLLVNGAGVIYNG 322
QY 297	PKLIGSGSNVQFYQTNVDIEIPDANQFEIVHSELNQNPK 336
DB 323	PKPLGQSGPMLRLDQ--LSLVTAEHVRITETSLVRDSVK 360
RESULT 14	
T12994	
riboflavin biosynthesis protein ribG homolog T21L8.140 - Arabidopsis thaliana	
C:Species: Arabidopsis thaliana (mouse-ear cress)	
C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 26-May-2000	
C:Accession: T12994	
R:Choisne, N.; Robert, C.; Brottier, P.; Wincker, P.; Cattolico, L.; Artiguenave, F.; submitted to the Protein Sequence Database, July 1999	
A:Reference number: Z17586	
A:Accession: T12994	
A:Molecule type: DNA	
A:Residues: 1-599 <CHO>	
A:Cross-references: EMBL:AL096860; GSPDB:GN00061; ATSP:T21L8.140	
A:Experimental source: cultivar Columbia; BAC clone T21L8	
C:Genetics:	
A:Gene: ATSP:T21L8.140	
A:Map position: 3	
A:Introns: 119/3; 164/3; 198/1; 240/1; 344/3; 384/3; 437/1; 513/3	
Query Match 17.9%; Score 320; DB 2; Length 599;	
Best Local Similarity 25.9%; Pred. No. 1.3e-16;	
Matches 91; Conservative 69; Mismatches 145; Indels 46; Gaps 8;	
QY 4	ATQLANVVGQTVNPPVCAVY-VNIEGRIVGICAHLRKGDHAEVQALDMAQNAEGATI 62
DB 39	ADLSEMSAGLTSPPHPCFCVITATSSGKVAGEGYLYAQGTKPAELAVAAAGEFSRGATA 98
QY 63	YITLPCSHFGSTPPCVNKKIIDCKIAKVYATKNSLDTHGD--ETLRAHGEVEECVDD 120
DB 99	YLNMEPGDCHGD-HTAVSALVQAGIERVVVGRHPHLQHLRGSARIELRSHGIEVNVLGED 157
QY 121	RASQLYQDFPKA-----KAKQLPQITVVKVASLDGKQANDNCSOWITNKEVKQ 169
DB 158	FESKVELEDARKSCLLVNAPLIHRACSRVPFVLKYAMTLDGKIAASGSHAAMISSKLSRT 217
QY 170	DYVKLRHRHDAVLGRRRTVELDDPQVTTTQDGKNPIKVILSKSGNHFNQOIYQDESTEP 229
DB 218	RVFELRGGSDAVIVGNTVRQDDPLRTARHGQHTPTTRIVMQSLDL-----PEKAN 269
QY 230	IWIYTNPNLTSTNQ-----THIEIYILKSCDLTTILHNLKRGVGTLLVEAG 276
DB 270	LNDVSEVSIIVVTQCGARKSFQKLLASKGVEVVEFDMLNPREVMEYFHLRGYLSILWECC 329
QY 277	PT-TTSEFSIYIIDEFILYAPKLIGGS-----GNVQFYQTNVDIEI 317

Db 330 GTLAASAISSSVIHKVVAFVAPKIIIGSKAPSPVGDGLGMVEMTQALNLIDV 380

## RESULT 15

Tl0638  
hypothetical protein Tl3K14.120 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 20-Apr-2000  
C:Accession: Tl0638  
R:Bevan, M.; Pohl, T.; Weizenegger, T.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.; Lemcke  
submitted to the Protein Sequence Database, June 1999  
A:Reference number: Z16991  
A:Accession: Tl0638  
A:Molecule type: DNA  
A:Residues: 1-363 <BEV>  
A:Cross-references: EMBL:AL080282; GSPDB:GN00062; ATSP:Tl3K14.120  
A:Experimental source: cultivar Columbia; BAC clone Tl3K14  
C:Genetics:  
A:Gene: ATSP:Tl3K14.120  
A:Map position: 4  
A:Introns: 60/3; 160/2  
C:Superfamily: Chlamydomophila pneumoniae riboflavin-specific deaminase

Query Match 16.88; Score 300; DB 2; Length 363;  
Best Local Similarity 28.48; Pred. No. 2.1e-15;  
Matches 83; Conservative 61; Mismatches 118; Indels 30; Gaps 8;  
QY 1 MDYAIQLANNVOGOTGVNPPGAVVNEGRIVGIGAHLRKGDHAEVQALDMAQOONAEA 60  
Db 15 MRKCVELAKRAIGCTSPNPWGCIVIVDGDIVGQGFHPKAGQPHAEVFAIRDAGELAENA 74  
QY 61 TIYITLPCSHFGSTPPCVNKIIDCKIAKVYATKDNS--LDTHGDETFLRAHGIEVE-CV 117  
Db 75 TAYVSLPCNHYGRTPPCTEALIKAKVRVVGWVDPNPIVFSSGISRLKDGADIDTVSV 134  
QY 118 DDERASQLYODFFKAKAKQLPQITVKVSASLD-----CKQANDNGCSQWITNKEVKQD 170  
Db 135 EEELCKRMNEGFTHRMLTGTGKPFLLALRYSMVNGCLLDKICGASDSG-----G 182  
QY 171 VY-KLRRHDA-VLTGRTVELDDPQVTTTIQDGNPKIKVIL---SKSGNTHFNQIYQD 225  
Db 183 YYSKLQYDAITLSSLSDELSSISQEAINVSIQPIQIIVASNAQOSHILASSHTVEE 242  
QY 226 ESTPIWIYTNPNLTS---NOTHIEIYLLKSCDLTTTLHNLKRGVGTLLVE 274  
Db 243 SGPKVVVFTAKESVAESGSSGVETVVLEKINLDSILDYCYNRGLCSVLLD 294

Search completed: November 14, 2000, 10:08:13  
Job time: 91 sec

**THIS PAGE BLANK (USPTO)**



```

CC CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE HTP
CC CC REDUCTASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L09228; AAA67481.1; -.
CC DR EMBL: X51510; CAA35878.1; -.
CC DR EMBL: Z99116; CAB14260.1; -.
CC DR PIR: A35711; A35711.
CC DR PIR: S45543; S45543.
CC DR SUBTILIST: BG10518; RIBD.
CC DR INTERPRO: IPR002125; -.
CC DR PFAM: PF01872; RibD_C; 1.
CC DR PFAM: PF00383; dCMP_cyt_deam; 1.
CC DR PROSITE: PS00903; CYT_DCMP_DEAMINASES; 1.
CC DR Riboflavin biosynthesis; Hydrolase; Zinc; Oxidoreductase; NADP;
CC KW Multifunctional enzyme.
CC FT DOMAIN 1 144 DEAMINASE.
CC FT METAL 145 361 REDUCTASE.
CC FT METAL 49 49 ZINC (BY SIMILARITY).
CC FT METAL 74 74 ZINC (BY SIMILARITY).
CC FT METAL 83 83 ZINC (BY SIMILARITY).
CC SQ SEQUENCE 361 AA; 39305 MW; DA836930BFDECA3B CRC64;

Query Match 30.6%; Score 547; DB 1; Length 361;
Best Local Similarity 36.1%; Pred. No. 9.1e-35;
Matches 130; Conservative 68; Mismatches 134; Indels 28; Gaps 8;

QY 1 MDAIOLANMVOGQTGVPVGAHVNVNNEGRVIGIGIAHLRKGDKHAEVQALDMAQQAEGA 60
Db 6 MKLALDAKQGEQTESPLVGAHVVDQGVGMGAHLKYGAEHAEVHAHMAHAEGA 65

QY 61 TYITILEPCSHFGSTPPCVNKKIIDCKIAKVYATKDNS--LDTHGDETIRAHGIEV-ECV 117
Db 66 DYIVTLEPCSHYKTPPCAEILINSIGIKRVFVAMRDPNPLVAGRGISMKEAGIEVREGI 125

QY 118 DDERASQYODFFKAKAKOLPOITVKVSASLDGKQANDGQOWITNKVKQDVYKLRHR 177
Db 126 LADQAEKLEKELHFHRTGLPVYTLKAAASLDGKIATSTGDSKWTSEARQDAQYRKT 185

QY 178 HDAVLTGRRTEVLDDPQYTRIQD-CGNPKTKVLSKSGNIHFNQIYODESTPIWIYT-- 234
Db 186 HQSILVGVGTVRADNPSLCRLPNVTQKQVRVILDTVLSIPEDAKVICQIAPTWIFTTA 245

QY 235 -----ENPNLTSTNQTHIEIYLKSCDLTILHNLKRGVGTLLVEAGPTTSEF-SIYI 288
Db 246 RADEEKRLKLSAFGVNIFTLETRIQIPDVLKILAEIGMSVVEGGSVAHGSFVKEGCF 305

QY 289 DEFILYAPKLIGG-----SG-NYQFYQTDNDVIPDANQFEIVHSELLNQNVKLTLR 340
Db 306 QEIIFFAPKLIGGTHAPSLISGEGFSQMKDVPQLQFTDITQ-----IGDKILTKAK 357

RESULT 2
RIBD.AQAE
ID RIBD.AQAE STANDARD; PRT; 356 AA.
AC 066534.
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE RIBOFLAVIN BIOSYNTHESIS PROTEIN RIBD [INCLUDES:
DE DIAMINOHYDROXYPHOSPHORIBOSYLAMINOPYRIMIDINE DEAMINASE (EC 3.5.4.26)
DE (RIBOFLAVIN-SPECIFIC DEAMINASE); 5-AMINO-6-(5-
DE PHOSPHORIBOSYLAMINO)URACIL REDUCTASE (EC 1.1.1.193) (HTP REDUCTASE)].
DE RIBD OR RIBG OR AQ_138.

```

```

OS Aquifex aeolicus.
CC Bacteria; Aquificales; Aquificaceae; Aquifex.
CC [1]
RN SEQUENCE FROM N.A.
CC STRAIN-VF5;
CC MEDLINE: 96196666.
CC Decker G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
CC Graham D.E., Overbeck R., Sneed M.A., Keller M., Aujay M., Huber R.,
CC Feldman R.A., Short J.M., Olson G.J., Swanson R.V.:
CC "The complete genome of the hyperthermophilic bacterium Aquifex
CC aeolicus."
CC Nature 392:353-358(1998).
CC RL Nature 392:353-358(1998).
CC CC -!- FUNCTION: CONVERTS 2,5-DIAMINO-6-(RIBOSYLAMINO)-4(3H)-PYRIMIDINONE
CC 5'-PHOSPHATE INTO 5-AMINO-6-(RIBOSYLAMINO)-2,4(1H,3H)-
CC PYRIMIDINEDIONE 5'-PHOSPHATE.
CC CC -!- CATALYTIC ACTIVITY: 2,5-DIAMINO-6-HYDROXY-4-(5-
CC PHOSPHORIBOSYLAMINO)PYRIMIDINE + H(2)O = 5-AMINO-6-(5-
CC PHOSPHORIBOSYLAMINO)URACIL + NH(3).
CC CC -!- CATALYTIC ACTIVITY: 5-AMINO-6-(5-PHOSPHORIBOSYLAMINO)URACIL +
CC NADP(+) = 5-AMINO-6-(5-PHOSPHORIBOSYLAMINO)URACIL + NADPH.
CC CC -!- COFACTOR: ZINC (BY SIMILARITY).
CC CC -!- PATHWAY: SECOND AN THIRD STEPS OF RIBOFLAVIN BIOSYNTHESIS.
CC CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE CYTIDINE AND
CC DEOXYCYTIDYLATE DEAMINASES FAMILY.
CC CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE HTP
CC REDUCTASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE000675; AAC06487.1; -.
CC DR INTERPRO: IPR002125; -.
CC DR INTERPRO: IPR002734; -.
CC DR PFAM: PF01872; RibD_C; 1.
CC DR PFAM: PF00383; dCMP_cyt_deam; 1.
CC DR PROSITE: PS00903; CYT_DCMP_DEAMINASES; 1.
CC DR Riboflavin biosynthesis; Hydrolase; Zinc; Oxidoreductase; NADP;
CC KW Multifunctional enzyme.
CC FT DOMAIN 1 148 DEAMINASE.
CC FT METAL 149 356 REDUCTASE.
CC FT METAL 53 53 ZINC (BY SIMILARITY).
CC FT METAL 78 78 ZINC (BY SIMILARITY).
CC FT METAL 87 87 ZINC (BY SIMILARITY).
CC SQ SEQUENCE 356 AA; 39628 MW; 7775515B1E4D02BD CRC64;

Query Match 30.5%; Score 545; DB 1; Length 356;
Best Local Similarity 40.5%; Pred. No. 1.3e-34;
Matches 125; Conservative 54; Mismatches 122; Indels 8; Gaps 5;

QY 1 MDAIOLANMVOGQTGVPVGAHVNVNNEGRVIGIGIAHLRKGDKHAEVQALDMAQQAEGA 60
Db 10 MKLALSLAKRKGYPHPNPTVGAVVKEGIVGLGHEKAGRPHAEVMAVGQAGEKAKGA 69

QY 61 TYITILEPCSHFGSTPPCVNKKIIDCKIAKVYATKDNS--LDTHGDETIRAHGIEV-ECV 117
Db 70 TLYVTLEPCTHFGRTPPCTDAIIRSGIKRVVATLDPNPLMSGKGVKELRNAGIEVDVG 129

QY 118 DDERASQYODFFKAKAKOLPOITVKVSASLDGKQANDGQOWITNKVKQDVYKLRHR 177
Db 130 CBEAEHLENEDEFTYITQRPVITLKAQTLGKALTLTGSSKWTSEKSRKVAHILRE 189

QY 178 HDAVLTGRRTEVLDDPQYTRIQD-CGNPKTKVLSKSGNIHFNQIYODESTPIWIYTEN 236
Db 190 ATAVLVGVNTVIKDDPHLTVRFVPTKEQVPIILDPELVEPLSAKVLNTEEAPTIVITKK 249

QY 237 PNLTSTNQTHIEIYLKSCDLTILHNLKRGVGTLLVEAGP-TTTSFSESIYIDEFI 292

```

```

Db 250 ENERAEKLEKGVOLILKGFNLKILKLEIMHLMVEGGPRTLTSLFKEGFFDRIV 309
Qy 293 LYPAPKLIG 301
Db 310 VFIAPKMG 318

RESULT 3
ID RIBD_BACAM STANDARD; PRT; 371 AA.
AC P70814;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE RIBOFLAVIN BIOSYNTHESIS PROTEIN RIBD [INCLUDES:
DE DIAMINOHYDROXYPHOSPHORIBOSYLAMINOPYRIMIDINE DEAMINASE (EC 3.5.4.26)
DE (RIBOFLAVIN-SPECIFIC DEAMINASE); 5-AMINO-6-(5-
DE PHOSPHORIBOSYLAMINO)URACIL REDUCTASE (EC 1.1.1.193) (HTP REDUCTASE)].
GN RIBD OR RIBG.
OS Bacillus amyloliquefaciens.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A 50.
RX MEDLINE; 97376523.
RA Gusarov I.I., Krenova R.A., Podcharniaev D.A., Iomantas I.U.V.,
RA Abalakiina E.G., Stoinova N.V., Perumov D.A., Kozlov I.U.I.;
RT "Riboflavin biosynthetic genes in Bacillus amyloliquefaciens: primary
RT structure, organization and regulation of activity.";
RL Mol. Biol. (Mosk) 31:446-453(1997).
CC -1- FUNCTION: CONVERTS 2,5-DIAMINO-6-(RIBOSYLAMINO)-4(3H)-PYRIMIDINONE
CC 5'-PHOSPHATE INTO 5'-AMINO-6-(RIBOSYLAMINO)-2,4(1H,3H)-
CC PYRIMIDINEDIONE 5'-PHOSPHATE.
CC -1- CATALYTIC ACTIVITY: 2,5-DIAMINO-6-HYDROXY-4-(5-
CC PHOSPHORIBOSYLAMINO)PYRIMIDINE + H(2)O = 5-AMINO-6-(5-
CC PHOSPHORIBOSYLAMINO)URACIL + NH(3).
CC -1- CATALYTIC ACTIVITY: 5-AMINO-6-(5-PHOSPHORIBOSYLAMINO)URACIL +
CC NADP(+) = 5-AMINO-6-(5-PHOSPHORIBOSYLAMINO)URACIL + NADPH.
CC -1- COFACTOR: ZINC (BY SIMILARITY).
CC -1- PATHWAY: SECOND AN THIRD STEPS OF RIBOFLAVIN BIOSYNTHESIS.
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE CYTIDINE AND
CC DEOXYCYTIDYLATE DEAMINASES FAMILY.
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE HTP
CC REDUCTASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@lsb-sib.ch).
CC
CC EMBL; X95955; CAA65189.1; -.
CC INTERPRO; IPR002125; -.
CC INTERPRO; IPR002734; -.
CC PRAM; PF01872; RibD_C; 1.
CC PRAM; PF00383; dCMP_cyt_deam; 1.
CC PROSITE; PS00903; Cyt_dCMP_deaminases; 1.
KW Riboflavin biosynthesis; Hydrolase; Zinc; Oxidoreductase; NADP;
KW Multifunctional enzyme.
FT DOMAIN 1 144 DEAMINASE.
FT METAL 145 371 REDUCTASE.
FT METAL 49 49 ZINC (BY SIMILARITY).
FT METAL 74 74 ZINC (BY SIMILARITY).
FT METAL 83 83 ZINC (BY SIMILARITY).
SQ SEQUENCE 371 AA; 40166 MW; 1A90722056A6BBA9 CRC64;

Query Match 28.7%; Score 513; DB 1; Length 371;
Best Local Similarity 34.9%; Pred. No. 3.8e-32;
Matches 122; Conservative 62; Mismatches 154; Indels 12; Gaps 5;

```

```

Qy 1 MDYAIQIANVVGQGVNPPVGVAVVNEGRIVGTGAHLRKGDKHAEOALDMAOONAEGA 60
Db 6 MNTAIELARRGEGQTPNPLVGVAVVVKRQIVGMGAHLQYGEAHEAHEAIAINMAGSLAKGA 65
Qy 61 TIXITLEPCSHFGSTPPCVNKNLIDCKIAKVYATKDNS--LDTHGDETLRAHGLEVEVC-V 117
Db 66 DLVITLEPCSHYGTTPCAELIMKSGIKRPIAIVEDNPLVAGKGTMLERAGLEVITGL 125
Qy 118 DDERASOLYQDFFKAKAKAQPOITVKVVSASLDGKQANDNGOSOWITNKEVKODVYKLRHR 177
Db 126 LRQAEELNKMFLHFMRTGLPYVTLKAASLDGKTATETGDSKWTSEARLDAQQVRS 185
Qy 178 HDAVLTRGRVVELDDPYTTRIQDG-KNPIKVLKSGNHFNOQIYODESTPIWITEN 236
Db 186 HQRLVAGTVKADNPISLTCRLPDVAKQPVRLDITKLTVPETANVLTDGAAPTWIFTAA 245
Qy 237 PNLTSNTHIEIIVLKSCDLTT-----ILHNLKRGVGTLLVEAGPTTISEF-SIYYI 288
Db 246 GSDVRKKDRLTALGIKVTLETDRHIPEVLSILAENGIMSVYEGGASVHGSEVKGACF 305
Qy 289 DEFILYAPKLIGGSGNYYQVQTNVDVIEIPDANQFETVHSELLMONVKLT 338
Db 306 DELHFYFAPILIGTGLAPSLISGEGFQSMKDVPHLQFTQITQIGDPDKLT 355

RESULT 4
ID RIBD_HAEIN STANDARD; PRT; 372 AA.
AC P44326;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE RIBOFLAVIN BIOSYNTHESIS PROTEIN RIBD [INCLUDES:
DE DIAMINOHYDROXYPHOSPHORIBOSYLAMINOPYRIMIDINE DEAMINASE (EC 3.5.4.26)
DE (RIBOFLAVIN-SPECIFIC DEAMINASE); 5-AMINO-6-(5-
DE PHOSPHORIBOSYLAMINO)URACIL REDUCTASE (EC 1.1.1.193) (HTP REDUCTASE)].
GN RIBD OR RIBG OR HI0944.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-RD / KW20;
RX MEDLINE; 95350630.
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus
RT influenzae Rd.";
RL Science 269:496-512(1995).
CC -1- FUNCTION: CONVERTS 2,5-DIAMINO-6-(RIBOSYLAMINO)-4(3H)-PYRIMIDINONE
CC 5'-PHOSPHATE INTO 5'-AMINO-6-(RIBOSYLAMINO)-2,4(1H,3H)-
CC PYRIMIDINEDIONE 5'-PHOSPHATE.
CC -1- CATALYTIC ACTIVITY: 2,5-DIAMINO-6-HYDROXY-4-(5-
CC PHOSPHORIBOSYLAMINO)PYRIMIDINE + H(2)O = 5-AMINO-6-(5-
CC PHOSPHORIBOSYLAMINO)URACIL + NH(3).
CC -1- CATALYTIC ACTIVITY: 5-AMINO-6-(5-PHOSPHORIBOSYLAMINO)URACIL +
CC NADP(+) = 5-AMINO-6-(5-PHOSPHORIBOSYLAMINO)URACIL + NADPH.
CC -1- COFACTOR: ZINC (BY SIMILARITY).
CC -1- PATHWAY: SECOND AN THIRD STEPS OF RIBOFLAVIN BIOSYNTHESIS.
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE CYTIDINE AND
CC DEOXYCYTIDYLATE DEAMINASES FAMILY.
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE HTP
CC REDUCTASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 DR EMBL; U32775; AAC22598.1; -;  
 DR TIGR; H10944; -;  
 DR INTERPRO: IPR002125; -;  
 DR INTERPRO: IPR002734; -;  
 DR PFAM; PF01872; RIBD\_C; 1.  
 DR PFAM; PF00383; dCMP\_cyt\_deam; 1.  
 DR PROSITE; PS00903; CYT\_DCMP\_DEAMINASES; 1.  
 KW Riboflavin biosynthesis; Hydrolase; Zinc; Oxidoreductase; NADP;  
 KW Multifunctional enzyme.  
 FT DOMAIN 1 150 DEAMINASE.  
 FT DOMAIN 151 372 REDUCTASE.  
 FT METAL 55 55 ZINC (BY SIMILARITY).  
 FT METAL 80 80 ZINC (BY SIMILARITY).  
 FT METAL 89 89 ZINC (BY SIMILARITY).  
 SQ SEQUENCE 372 AA; 40931 MW; C7504EDD8B86F7D CRC64;

Query Match 27.7%; Score 495.5; DB 1; Length 372;  
 Best Local Similarity 34.3%; Pred. No. 8.2e-31;  
 Matches 124; Conservative 67; Mismatches 148; Indels 23; Gaps 9;

QY 1 MDYAIQLANNVQGTGVPVAVVYNEGRVIGVIGAHLRKGDHAEVQALDMAQNAEGA 60  
 DB 12 MQRALDLAAKQGYTTTPNPVSGCVLVKNGEIVGEGFHFKAGQPHAEVVALAQAGENAKGA 71  
 QY 61 TIVITLPCSHFGSTPPCVNKIIDCKIAKVYATKNSLDTHGCD--ETLRAHGIE--VEC 116  
 DB 72 TAVITLPCAHYGRTPPCALGLIEAGVVKVIAAMQDPNPQVAGKGKGLKMLSDAGIESTVNL 131  
 QY 117 VDDEASQLYQDFKAKAKOLPQITVYKVSASLDGKQANDNGSQWITNKREVKODVYKLRH 176  
 DB 132 LNDQ-AEKINGKFLKMRQMGPPVQLKSLAMDGRMTAMAGESKWTGPDARSVDQKMR 190  
 QY 177 RHDAVLTRGVETLDDPOYTRIQDG-----KNPIKVLKSGNHFNQOIY 223  
 DB 191 KSSALLSTVTIADPDSLVNVRWDEFPENLKTEYKKEWLQPVVRVLDLSOHRIQPTHKLF 250  
 QY 224 QDSTPIWIYTNP-NLTSNQTHIE-IIVLKSCDLTLHLNLYKRGVTLVVEAGPTT- 280  
 DB 251 LTHS-PWLVSSPRDLTGPPDFCEQIIFPKENLLKELRELKRGQINTLWVEAGANLSC 309  
 QY 281 SEFSYIYIDEFILIYAPKLIGSGNYQFYOTNDVIEIPDANOFIEVHSELLNONVKLTLR 340  
 DB 310 SLIDAKLVDELIIYIAPKLIGDNAR-GLCQLPNLTKLADAPLWQLNELEIGDIDKLTYT 368  
 QY 341 KK 342  
 DB 369 PK 370

RESULT 5  
 RIBD\_ECOLI  
 ID RIBD\_ECOLI STANDARD; PRT; 367 AA.  
 AC P25539;  
 DT 01-MAY-1992 (Rel. 22, Created)  
 DT 01-MAY-1992 (Rel. 22, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE RIBOFLAVIN BIOSYNTHESIS PROTEIN RIBD [INCLUDES:  
 DE DIAMINOHYDROXYPHOSPHORIBOSYLAMINOPYRIMIDINE DEAMINASE (EC 3.5.4.26)  
 DE (RIBOFLAVIN-SPECIFIC DEAMINASE); 5-AMINO-6-(5-  
 DE PHOSPHORIBOSYLAMINO)URACIL REDUCTASE (EC 1.1.1.193) (HTP REDUCTASE)].  
 GN RIBD OR RIBG.  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 RN [1]

RP SEQUENCE FROM N.A.  
 RX STRAIN-K12;  
 RX MEDLINE; 93024316.  
 RA Taura T., Ueguchi C., Shiba K., Ito K.;  
 RT "Insertional disruption of the nusB (ssyb) gene leads to cold-  
 RT sensitive growth of Escherichia coli and suppression of the secY24  
 RT mutation.";  
 RT Mol. Gen. Genet. 234:429-432(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN-K12 / MG1655;  
 RX MEDLINE; 97426617.  
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
 RA Mau B., Shao Y.;  
 RT "The complete genome sequence of Escherichia coli K-12.";  
 RT Science 277:1453-1474(1997).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Roberts D., Allen E., Araujo R., Aparicio A., Chung E., Davis K.,  
 RA Duncan M., Federspiel N., Hyman R., Kalman S., Komp C., Kurdi O.,  
 RA Lew H., Lin D., Mamath A., Oefner P., Schramm S., Davis R.W.;  
 RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP CHARACTERIZATION.  
 RX MEDLINE; 97221604.  
 RA Richter G., Fischer M., Krieger C., Eberhardt S., Luttgen H.,  
 RA Gerstenschlager I., Bacher A.;  
 RT "Biosynthesis of riboflavin: characterization of the bifunctional  
 RT deaminase-reductase of Escherichia coli and Bacillus subtilis.";  
 RL J. Bacteriol. 179:2022-2028(1997).  
 CC -!- FUNCTION: CONVERTS 2,5-DIAMINO-6-(RIBOSYLAMINO)-4(3H)-PYRIMIDINONE  
 CC 5'-PHOSPHATE INTO 5-AMINO-6-(RIBOSYLAMINO)-2,4(1H,3H)-  
 CC PYRIMIDINEDIONE 5'-PHOSPHATE.  
 CC -!- CATALYTIC ACTIVITY: 2,5-DIAMINO-6-HYDROXY-4-(5-  
 CC PHOSPHORIBOSYLAMINO)PYRIMIDINE + H(2)O = 5-AMINO-6-(5-  
 CC PHOSPHORIBOSYLAMINO)URACIL + NH(3).  
 CC -!- CATALYTIC ACTIVITY: 5-AMINO-6-(5-PHOSPHORIBITYLAMINO)URACIL +  
 CC NADP(+) = 5-AMINO-6-(5-PHOSPHORIBOSYLAMINO)URACIL + NADPH.  
 CC -!- COFACTOR: ZINC (BY SIMILARITY).  
 CC -!- PATHWAY: SECOND AN THIRD STEPS OF RIBOFLAVIN BIOSYNTHESIS.  
 CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE CYTIDINE AND  
 CC DEOXYCYTIDYLATE DEAMINASES FAMILY.  
 CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE HTP  
 CC REDUCTASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 DR EMBL; X64395; CAA45735.1; -;  
 DR EMBL; AE000148; AAC73517.1; -;  
 DR EMBL; U82664; AAB40170.1; -;  
 DR PIR; S26201; S26201.  
 DR ECOGENE; EG11321; RIBD.  
 DR INTERPRO; IPR002125; -;  
 DR INTERPRO; IPR002734; -;  
 DR PFAM; PF01872; RIBD\_C; 1.  
 DR PFAM; PF00383; dCMP\_cyt\_deam; 1.  
 DR PROSITE; PS00903; CYT\_DCMP\_DEAMINASES; 1.  
 KW Riboflavin biosynthesis; Hydrolase; Zinc; Oxidoreductase; NADP;  
 KW Multifunctional enzyme.  
 FT DOMAIN 1 145 DEAMINASE.  
 FT DOMAIN 146 367 REDUCTASE.  
 FT METAL 50 50 ZINC (BY SIMILARITY).  
 FT METAL 75 75 ZINC (BY SIMILARITY).  
 FT METAL 84 84 ZINC (BY SIMILARITY).  
 SQ SEQUENCE 367 AA; 40338 MW; B19CEFA74D48D14D CRC64;





RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,  
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,  
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
RA Rutter S., Seeger K., Skelton S., Squares S., Squires R., Sulston J.E.,  
RA Taylor K., Whitehead S., Barrell B.G.:  
RT "Deciphering the biology of Mycobacterium tuberculosis from the  
RT complete genome sequence.";  
RL Nature 393:537-544(1998).  
CC -!- FUNCTION: CONVERTS 2,5-DIAMINO-6-(RIBOSYLAMINO)-4(3H)-PYRIMIDINONE  
CC 5'-PHOSPHATE INTO 5-AMINO-6-(RIBOSYLAMINO)-2,4(1H,3H)-  
CC PYRIMIDINEDIONE 5'-PHOSPHATE.  
CC -!- CATALYTIC ACTIVITY: 2,5-DIAMINO-6-HYDROXY-4-(5-  
CC PHOSPHORIBOSYLAMINO)PYRIMIDINE + H(2)O = 5-AMINO-6-(5-  
CC PHOSPHORIBOSYLAMINO)URACIL + NH(3).  
CC -!- CATALYTIC ACTIVITY: 5-AMINO-6-(5-PHOSPHORIBOSYLAMINO)URACIL +  
CC NADP(+) = 5-AMINO-6-(5-PHOSPHORIBOSYLAMINO)URACIL + NADPH.  
CC -!- COFACTOR: ZINC (BY SIMILARITY).  
CC -!- PATHWAY: SECOND AN THIRD STEPS OF RIBOFLAVIN BIOSYNTHESIS.  
CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE CYTIDINE AND  
CC DEOXYCYTIDYLATE DEAMINASES FAMILY.  
CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE HTP  
CC REDUCTASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; 280108; CAB02188.1; -;  
DR TUBERCULIST; RV1409; -;  
DR INTERPRO; IPR002125; -;  
DR INTERPRO; IPR002734; -;  
DR PFAM; PF01872; RIBD\_C; 1.  
DR PFAM; PF00383; dCMP\_cyt\_deam; 1.  
DR PROSITE; PS00903; CYT\_DCMP\_DEAMINASES; 1.  
KW Riboflavin biosynthesis; Hydroxylase; Zinc; Oxidoreductase; NADP;  
KW Multifunctional enzyme.  
FT DOMAIN 1 152 DEAMINASE.  
FT METAL 53 339 REDUCTASE.  
FT METAL 57 57 ZINC (BY SIMILARITY).  
FT METAL 82 82 ZINC (BY SIMILARITY).  
FT METAL 91 91 ZINC (BY SIMILARITY).  
SQ SEQUENCE 339 AA; 35366 MW; 1DC0A1B8E7EC08B4 CRC64;  
-----  
Query Match 24.3%; Score 434.5; DB 1; Length 339;  
Best Local Similarity 35.0%; Pred. No. 3.3e-26;  
Matches 109; Conservative 57; Mismatches 114; Indels 31; Gaps 9;  
-----  
QY 1 MDYATQLANNVQGTGVNPPVGVAVVNV-EGRIYIGAHILKRGDKHAEVQALDMAQNAEG 59  
DB 13 MGLAIEHSYQVGGTKTPPKPPGVAVVDNPNRIVGAGTEPAGGDHAEVVALRRAGSLAG 72  
-----  
QY 60 ATIVITLEPCSHRGSPPCVKNK1IDCKIAKVYATKD-NSLDTHGDETLRHAHGIEVE--C 116  
DB 73 AIVVWTFEPCNHVGTTPPCVNAIEARVGTIVYAVADPNAGIAGGAGRLSAGLQVRSVG 132  
-----  
QY 117 VDDERASQLYQDFFKAKAKOLPOITVYKVSASLDGKQANDNGSQWITNKVEVKDYKLRH 176  
DB 133 LAEQVAAGPLREWLHKQRTGLPHVTWKYATSIDGRSAADGSSQWISSEARLDLHRRRA 192  
-----  
QY 177 RHDAVLTGRFTVLDLPQYTRTQDG-----KNPIKIVLSKSGNIHFNQIQDESTPIWI 232  
DB 193 IADAILVGTGTVLADDPALFARLADGSLAPQQLRVVWCKR-DIPPEARVLNDEARTMI 251  
-----  
QY 233 YTEPNPLTNSQTHIEIYILKSCDLTILHNLKRGVGTLLVEAGPTTSTSEF-SIYVIDEF 291  
DB 252 RTHEP-----MEVLRALS-DRTDV-----LLEGPTLAGAFLRAGAINRI 290

QY 292 ILYYAPKLIIGG 302  
DB 291 LAYVAPILLGG 301  
-----  
RESULT 8  
RIBD\_SYNY3 STANDARD; PRT; 368 AA.  
AC Q55158;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE RIBOFLAVIN BIOSYNTHESIS PROTEIN RIBD [INCLUDES:  
DE DIAMINOHYDROXYPHOSPHORIBOSYLAMINOPYRIMIDINE DEAMINASE (EC 3.5.4.26)  
DE (RIBOFLAVIN-SPECIFIC DEAMINASE); 5-AMINO-6-(5-  
DE PHOSPHORIBOSYLAMINO)URACIL REDUCTASE (EC 1.1.1.193) (HTP REDUCTASE)].  
GN RIBD OR SLR0066.  
OS Synechocystis sp. (strain PCC 6803).  
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 96127529.  
RA Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,  
RA Sugita M., Tabata S.:  
RT "Sequence analysis of the genome of the unicellular cyanobacterium  
RT Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb  
RT region from map positions 64% to 92% of the genome.";  
RL DNA Res. 2:153-166(1995).  
CC -!- FUNCTION: CONVERTS 2,5-DIAMINO-6-(RIBOSYLAMINO)-4(3H)-PYRIMIDINONE  
CC 5'-PHOSPHATE INTO 5-AMINO-6-(RIBOSYLAMINO)-2,4(1H,3H)-  
CC PYRIMIDINEDIONE 5'-PHOSPHATE.  
CC -!- CATALYTIC ACTIVITY: 2,5-DIAMINO-6-HYDROXY-4-(5-  
CC PHOSPHORIBOSYLAMINO)PYRIMIDINE + H(2)O = 5-AMINO-6-(5-  
CC PHOSPHORIBOSYLAMINO)URACIL + NH(3).  
CC -!- CATALYTIC ACTIVITY: 5-AMINO-6-(5-PHOSPHORIBOSYLAMINO)URACIL +  
CC NADP(+) = 5-AMINO-6-(5-PHOSPHORIBOSYLAMINO)URACIL + NADPH.  
CC -!- COFACTOR: ZINC (BY SIMILARITY).  
CC -!- PATHWAY: SECOND AN THIRD STEPS OF RIBOFLAVIN BIOSYNTHESIS.  
CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE CYTIDINE AND  
CC DEOXYCYTIDYLATE DEAMINASES FAMILY.  
CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE HTP  
CC REDUCTASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; D64001; BAA10295.1; -;  
DR INTERPRO; IPR002125; -;  
DR INTERPRO; IPR002734; -;  
DR PFAM; PF01872; RIBD\_C; 1.  
DR PFAM; PF00383; dCMP\_cyt\_deam; 1.  
DR PROSITE; PS00903; CYT\_DCMP\_DEAMINASES; 1.  
KW Riboflavin biosynthesis; Hydroxylase; Zinc; Oxidoreductase; NADP;  
KW Multifunctional enzyme.  
FT DOMAIN 1 146 DEAMINASE.  
FT METAL 51 368 REDUCTASE.  
FT METAL 76 76 ZINC (BY SIMILARITY).  
FT METAL 85 85 ZINC (BY SIMILARITY).  
SQ SEQUENCE 368 AA; 39995 MW; BCD8A8ED0916B50BC CRC64;  
-----  
Query Match 23.1%; Score 413.5; DB 1; Length 368;  
Best Local Similarity 32.3%; Pred. No. 1.5e-24;  
Matches 108; Conservative 54; Mismatches 157; Indels 15; Gaps 5;  
-----  
QY 1 MDYATQLANNVQGTGVNPPVGVAVVNVGRIYIGAHILKRGDKHAEVQALDMAQNAEGA 60



```
CC DEOXYCYTIDYLATE DEAMINASES FAMILY.
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE HTP
CC REDUCTASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB003693; BAA20053.1; -.
DR INTERPRO; IPR002125; -.
DR INTERPRO; IPR002734; -.
DR PFAM; PF01872; RibD_C; 1.
DR PFAM; PF00383; dCMP_cyt_deam; 1.
DR PROSITE; PS00903; CYT_DCMP_DEAMINASES; 1.
DR Riboflavin biosynthesis; Hydrolase; Zinc; Oxidoreductase; NADP;
KW Multifunctional enzyme.
FT DOMAIN 1 157 DEAMINASE.
FT DOMAIN 158 337 REDUCTASE.
FT METAL 63 63 ZINC (BY SIMILARITY).
FT METAL 88 88 ZINC (BY SIMILARITY).
FT METAL 97 97 ZINC (BY SIMILARITY).
SQ SEQUENCE 337 AA; 34888 MW; 2BBD4788BF2C243D CRC64;

Query Match 19.5%; Score 348; DB 1; Length 337;
Best Local Similarity 31.5%; Pred. No. 1.3e-19;
Matches 98; Conservative 38; Mismatches 123; Indels 52; Gaps 7;

QY 11 VQGTGVPVPGAVVNVN-EGRIVGIGAHRLKRGDKHAQVQALDMAQQNAEGATIIYITLPC 69
1:1:|||||:::|||||:|||||:|||||:|||||:|||||:
DQ 29 VRGTTSPVPGAVIIISTGSEIVGTGTPQVGGVHAQVQALADAAGTEGATAVVITLPC 88
1:1:|||||:::|||||:|||||:|||||:|||||:|||||:
QY 70 SHFGSTPPCVNKKIDCKIAKVATKD-NSLDTHGDDTLRAHGIE-VECVDDDERASOLYQ 127
1:1:|||||:::|||||:|||||:|||||:|||||:|||||:
DQ 89 RHTGRTGCTQALTEAGIKDVLFLHSDPNPNSAGGEGQVLVDAGINVVQLPSPGVPDALI 148
1:1:|||||:::|||||:|||||:|||||:|||||:|||||:
QY 128 DFFKAKAKQLPQITVKVSASLDGKQANDNGSQWITNKVKQDVYKLRHRHDAVLTRRT 187
1:1:|||||:::|||||:|||||:|||||:|||||:|||||:
DQ 149 PWLKSQVLRPHVTLKFAQTIDGFTAAADGTSOWITGEMARDVYHADREHRAIIIGTGT 208
1:1:|||||:::|||||:|||||:|||||:|||||:|||||:
QY 188 VELDDPQYTRIDQG-----KNPIKVL-----SKSGNIHFNOQIYQDESTPIW 231
1:1:|||||:::|||||:|||||:|||||:|||||:|||||:
DQ 209 ALIDNPSLTARYPDGTQREHQPRVVGIRNRNIADAGDAASNLNLGFEQATIDEA---- 264
1:1:|||||:::|||||:|||||:|||||:|||||:|||||:
QY 232 IYENPNLTSNQPHIEIYILKSCDLTTLNLNLYKRGVGTLLVRAGPTTTFSEFSIY-YIDE 290
1:1:|||||:::|||||:|||||:|||||:|||||:|||||:
DQ 265 -----LAELYATGARDVLVEGGAGLAGSFANQGLVDW 296
1:1:|||||:::|||||:|||||:|||||:|||||:|||||:
QY 291 FILYAPKLIG 301
1:1:|||||:::|||||:|||||:|||||:|||||:|||||:
DQ 297 LQVYQAPLLG 307
1:1:|||||:::|||||:|||||:|||||:|||||:|||||:

RESULT 11
RIB7_ARCFU
ID RIB7_ARCFU STANDARD; PRT; 219 AA.
AC O28272;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE PUTATIVE 5-AMINO-6-(5-PHOSPHORIBOSYLAMINO)URACIL REDUCTASE
DE (EC 1.1.1.193) (HTP REDUCTASE).
GN AF2007.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
OC Archaeoglobus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
```

```
RX MEDLINE; 98049343.
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Winn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,
RA Fleischnann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Coyne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artlich P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus."
RL Nature 390:364-370(1997).
CC -1- CATALYTIC ACTIVITY: 5-AMINO-6-(5-PHOSPHORIBOSYLAMINO)URACIL +
CC NADP(+) = 5-AMINO-6-(5-PHOSPHORIBOSYLAMINO)URACIL + NADPH.
CC -1- PATHWAY: THIRD STEP OF RIBOFLAVIN BIOSYNTHESIS.
CC -1- SIMILARITY: BELONGS TO THE HTP REDUCTASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000964; AAB89247.1; -.
DR TIGR; AF2007; -.
DR INTERPRO; IPR002734; -.
DR PFAM; PF01872; RibD_C; 1.
KW Hypothetical protein; Riboflavin biosynthesis; Oxidoreductase; NADP.
SQ SEQUENCE 219 AA; 24578 MW; 30DEA579024CF625 CRC64;

Query Match 10.9%; Score 195; DB 1; Length 219;
Best Local Similarity 29.8%; Pred. No. 3.7e-08;
Matches 54; Conservative 36; Mismatches 75; Indels 16; Gaps 4;

QY 138 PQITVKVSASLDGKQANDNGSQWITNKVKQDVYKLRHRHDAVLTRRTVELDDPQYTT 197
1:1:|||||:::|||||:|||||:|||||:|||||:|||||:
DQ 3 PYVFWVAASLDGKIDSERKQRIISCEEDLRIVDLRAESDAIMVGIGTVLADDPRLTV 62
1:1:|||||:::|||||:|||||:|||||:|||||:|||||:
QY 198 -----RIQDGK--NPVKVILSKSGNIHFNOQIYQDESTPIWYI-----ENPNLTSN 242
1:1:|||||:::|||||:|||||:|||||:|||||:|||||:
DQ 63 KSALREKQKQDKGKNPLRVVVDSCRVPPLTARILNDEARTLVASRTAPEKRVK 122
1:1:|||||:::|||||:|||||:|||||:|||||:|||||:
QY 243 QTHIEIYILKSCDLTTLNLNLYKRGVGTLLVRAGPTTTFSE-FSIYVIDEFLYAPKLIG 301
1:1:|||||:::|||||:|||||:|||||:|||||:|||||:
DQ 123 VAEVAVGEERVELSALLEFLHRKGVRRMLVEGGGTLISSLISSQLNVLDEIRIYYPFIG 182
1:1:|||||:::|||||:|||||:|||||:|||||:|||||:
QY 302 G 302
1:1:|||||:::|||||:|||||:|||||:|||||:|||||:
DQ 183 G 183

RESULT 12
RIB7_METJA
ID RIB7_METJA STANDARD; PRT; 224 AA.
AC Q58085;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE PUTATIVE 5-AMINO-6-(5-PHOSPHORIBOSYLAMINO)URACIL REDUCTASE
DE (EC 1.1.1.193) (HTP REDUCTASE).
GN MJ0671.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
OC Methanococcus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
```

```

RX MEDLINE; 96337999.
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kervilange A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geobach N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Uterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RL jannaschii."
RL Science 273:1058-1073(1996).
CC -!- CATALYTIC ACTIVITY: 5-AMINO-6-(5-PHOSPHORIBITYLAMINO)URACIL +
CC NADP(+) = 5-AMINO-6-(5-PHOSPHORIBOSYLAMINO)URACIL + NADPH.
CC -!- PATHWAY: THIRD STEP OF RIBOFLAVIN BIOSYNTHESIS.
CC -!- SIMILARITY: BELONGS TO THE HTP REDUCTASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U67514; AAB98665.1; -.
DR TIGR; MJ0671; -.
DR INTERPRO: IPR002734; -.
DR PFAM: PF01872; RibD_C; 1.
KW Hypothetical protein; Riboflavin biosynthesis; Oxidoreductase; NADP.
SQ SEQUENCE 234 AA; 25037 MW; 4D8C15CE291E89D8 CRC64;

Query Match 10.7%; Score 191; DB 1; Length 224;
Best Local Similarity 28.0%; Pred. No. 7.8e-08;
Matches 61; Conservative 41; Mismatches 82; Indels 34; Gaps 9;

Qy 138 POITVKVSASLDGKANDNGOSQWITNKVKQDVYKLRHRHDAVLGTRRTVELDDPQVTT 197
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Dy 9 PYIISNMGWTLGKLTATNDSR-ISCEEDLIRVHKIRANVDGIMVGIGTVLKDDPRLTV 67
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Qy 198 -RIQDKNPVKILSKSGNIHFNQIYODESTPIWIYTENPNLTSNTHIEI-----IYL 251
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Dy 68 HKISDRNPRIVVDSKURVPLNARVLNKKDAITATDTN-EEKEKKIKILEDGMGEV 126
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Qy 252 KSC-----DLTILNLYKRGVGLLVEAGPTTT-SEFSIYIIDEFILIYAPKILGGSGN 305
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Dy 127 VKCGRGKVDLKKMLDILYDKIKSILLEGCGTLNMGFKGLVDEVSVIAPKIFGCK-- 184
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Qy 306 YOFYOTNDVIEP---DANQFEIVHSELLNQNVKLTLR 340
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Dy 185 -----EAPTYVDGEGFKTV-----DECVKLEK 207
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |

RESULT 13
RIB7_METTH
ID RIB7_METTH STANDARD; PRT; 216 AA.
AC O26337;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 01-OCT-2000 (Rel. 39, Last annotation update)
DE PUTATIVE 5-AMINO-6-(5-PHOSPHORIBOSYLAMINO)URACIL REDUCTASE
DE (EC 1.1.1.193) (HTP REDUCTASE).
GN MTH235.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;
OC Methanobacterium.
CC [1]
RC STRAIN=DELTA H;
RP SEQUENCE FROM N.A.
RX SMITH-D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
RX MEDLINE; 98037514.
RA Smith D.R., Bashirzadeh R., Blakely D., Cook K., Gilbert K.,

```

```

RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,
RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA Jiwan N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Nolling J., Reeve J.N.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
RT delat: functional analysis and comparative genomics."
RL J. Bacteriol. 179:7135-7155(1997).
CC -!- CATALYTIC ACTIVITY: 5-AMINO-6-(5-PHOSPHORIBITYLAMINO)URACIL +
CC NADP(+) = 5-AMINO-6-(5-PHOSPHORIBOSYLAMINO)URACIL + NADPH.
CC -!- PATHWAY: THIRD STEP OF RIBOFLAVIN BIOSYNTHESIS.
CC -!- SIMILARITY: BELONGS TO THE HTP REDUCTASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000810; AAB84741.1; -.
DR INTERPRO: IPR002734; -.
DR PFAM: PF01872; RibD_C; 1.
KW Hypothetical protein; Riboflavin biosynthesis; Oxidoreductase; NADP.
SQ SEQUENCE 216 AA; 23576 MW; 436561C089DF41A6 CRC64;

Query Match 10.3%; Score 184; DB 1; Length 216;
Best Local Similarity 25.6%; Pred. No. 2.5e-07;
Matches 53; Conservative 46; Mismatches 98; Indels 10; Gaps 5;

Qy 138 POITVKVSASLDGKANDNGOSQWITNKVKQDVYKLRHRHDAVLGTRRTVELDDPQVTT 197
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Dy 3 PYVILNAAWTLGKLTATGSGE-ISGEEDLRRVHLELRCDALWINTVLADPRLTV 61
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Qy 198 RIQD---GKNPKIVILSKSGNIHFNQIYODESTPIWIYTEN---PNLTS---NTHIEI 249
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Dy 62 HRVDAAPGDNPRVVDSMARTPPHFRVLNDEAPTIVGSESAAPERVAELKRAEVVA 121
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Qy 250 YLKSCDLTILNLYKRGVGLLVEAGPTTT-SEFSIYIIDEFILIYAPKILGGSGNQV 308
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Dy 122 GTRRVLDLHLLERLHGMGIERLMLEGGSTLNSMLTGGVLDEVRCVAPMIVGGRDRTL 181
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Qy 309 YOTNDVIEIPDANQFEIVHSELLNQNV 335
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Dy 182 VDGEIDEMADAIRLEKRSYTLGEDL 208
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |

RESULT 14
RIB7_SULSO
ID RIB7_SULSO STANDARD; PRT; 213 AA.
AC P95872;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE PUTATIVE 5-AMINO-6-(5-PHOSPHORIBOSYLAMINO)URACIL REDUCTASE
DE (EC 1.1.1.193) (HTP REDUCTASE).
GN C06025.
OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
CC [1]
RC STRAIN=DSM 1617 / P2;
RX MEDLINE; 97055432.
RA Sensen C.W., Klenk H.-P., Singh R.K., Allard G., Chan C.C.-Y.,
RA Liu Q.Y., Penny S.L., Young F., Schenk M.E., Gaasterland T.,
RA Doolittle W.F., Ragan M.A., Charlebois R.L.;
RT "Organizational characteristics and information content of an
RT archaeal genome: 156 kb of sequence from Sulfolobus solfataricus
RT P2."
RL Mol. Microbiol. 22:175-191(1996).
CC -!- CATALYTIC ACTIVITY: 5-AMINO-6-(5-PHOSPHORIBITYLAMINO)URACIL +

```

```
CC      NADP(+) = 5-AMINO-6-(5-PHOSPHORIBOSYLAMINO)URACIL + NADPH.
CC      -1- PATHWAY: THIRD STEP OF RIBOFLAVIN BIOSYNTHESIS.
CC      -1- SIMILARITY: BELONGS TO THE HTP REDUCTASE FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL: Y08256; CAA69508.1; -.
CC      INTERPRO: IPR002734; -.
CC      PFAM: PF01872; RibD_C; 1.
CC      KW      Hypothetical protein; Riboflavin biosynthesis; Oxidoreductase; NADP.
CC      SEQUENCE 213 AA; 24561 MW; A741F6CADEF7631C CRC64;
CC      -----
Query Match          9.1%; Score 163; DB 1; Length 213;
Best Local Similarity 26.8%; Pred. No. 1e-05;
Matches 57; Conservative 48; Mismatches 92; Indels 16; Gaps 10;
CC      YQ 138 PQITVKVSASLDKQANDNCOSQWITNKVKQDVYKLRHRDAVLTRGRVVELDDPQYTT 197
CC      | : : : | : | : | : : | : | : | : | : | : | : | : | : | : | : |
CC      5 PYVIFSTVSIDGLATKTKGYSE-LSCPDKQKQHEIRSEVDAMVVGANTVRVDNPSLTV 63
CC      YQ 198 RI-QDGNPKIVILSKSGNIHFNOQIODESTPIWIITNPNTLSNOTHTIE--IIYLKSC 254
CC      | : : : | : | : | : : | : | : | : | : | : | : | : | : | : | : |
CC      64 KYGKRRNRPVRVTRSFNDPSYKIFITTPPSIV-IYTSYSEKVEEFKGVIVRKFL 122
CC      YQ 255 DLTTILHNLKXR-GVGFTLLVEAGPTTTFSEF---SIYVIDEFILYAPKLGGSGNYQFYQ 310
CC      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
CC      123 HLDDLEDLDNFRNVRRLWVEGGHLIWMFIKDNLY--DEIRITISPRIFG--NGVSFTQ 178
CC      YQ 311 TNDVIEIPDANQFEIVHSEL--LNQNVKFLRK 341
CC      | : : : | : | : | : : | : | : | : | : | : | : | : | : | : | : |
CC      179 GEGFIG-EDSPRIELIDAKICEGCEVHLYTKK 210
```

## RESULT 15

```
YFHC_ECOLI
ID      YFHC_ECOLI      STANDARD;      PRT;      178 AA.
AC      F30134;
DT      01-APR-1993 (Rel. 25, Created)
DT      01-APR-1993 (Rel. 25, Last sequence update)
DT      01-NOV-1997 (Rel. 35, Last annotation update)
DE      HYPOTHETICAL 20.0 KDA PROTEIN IN PURL-DPU INTERGENIC REGION (ORF178).
GN      YFHC.
OS      Escherichia coli.
OC      Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC      Escherichia.
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=NWL37;
RX      MEDLINE; 92292954.
RA      Poulsen L.K., Larsen N.W., Molin S., Andersson P.;
RT      "Analysis of an Escherichia coli mutant strain resistant to the cell-
RT      killing function encoded by the gef gene family.";
RL      Mol. Microbiol. 6:895-905(1992).
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=K12;
RA      Nashimoto H., Saito N.;
RL      Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.
RN      [3]
RP      SEQUENCE FROM N.A.
RC      STRAIN=K12 / MG1655;
RX      MEDLINE; 97426617.
RA      Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA      Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA      Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA      Mau B., Shao Y.;
```

```
RT      The complete genome sequence of Escherichia coli K-12.;
RL      Science 277:1453-1474(1997).
CC      -1- FUNCTION: MUTATION IN THIS PROTEIN MAKES E.COLI RESISTANT TO THE
CC      TOXIC PROTEINS ENCODED BY THE GEF GENE FAMILY.
CC      -1- COFACTOR: ZINC (BY SIMILARITY).
CC      -1- SIMILARITY: BELONGS TO THE CYTIDINE AND DEOXYCYTIDYLATE DEAMINASES
CC      FAMILY. STRONG, TO E.COLI AND B.SUBTILIS RIBG. STRONG, TO
CC      H.INFLUENZAE HI0906.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL: X72336; CAA51064.1; -.
CC      EMBL: D64044; BAA10909.1; -.
CC      EMBL: U36841; AAA79821.1; -.
CC      DR      EMBL: AE000342; AAC75612.1; -.
CC      PIR: S20974; S20974.
CC      ECOGENE: EG11372; YFHC.
CC      INTERPRO: IPR002125; -.
CC      DR      PFAM: PF00383; DCMPT_Cyt_deam; 1.
CC      DR      PROSITE: PS00903; CYT_DCMPT_DEAMINASES; 1.
CC      KW      Hypothetical protein; Hydrolase; zinc.
CC      FT      METAL 68 68 ZINC (BY SIMILARITY).
CC      FT      METAL 98 98 ZINC (BY SIMILARITY).
CC      FT      METAL 101 101 ZINC (BY SIMILARITY).
CC      FT      MUTAGEN 64 64 D->E: RESISTANCE TO THE CELL-KILLING
CC      FUNCTION ENCODED BY THE GEF GENE FAMILY.
CC      SEQUENCE 178 AA; 20026 MW; 80B2E3B5FD61AA8A CRC64;
CC      -----
Query Match          8.2%; Score 146.5; DB 1; Length 178;
Best Local Similarity 29.3%; Pred. No. 0.00014;
Matches 49; Conservative 29; Mismatches 60; Indels 29; Gaps 10;
CC      YQ 1 MDYAIOLANNVGQGTGVNPPGVAVVNEGRIVG-----IGAHLRKGDKHAQVAL--D 51
CC      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
CC      23 MRHALTLAKRWDEREV--PVGAVLVHNNRVIGEGWNRPIGRH--DPTAHAEIMALRQGG 78
CC      YQ 52 MAQQNAE--GATITITLEPCSHFGSTPPCVNKKIIDCKIAKVYVATKDNSLDTHGD--ETL 107
CC      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
CC      79 LVNQVRLIDATLYVTLEPC-----VMCAGAMHSRIGRVVFGARDAKTGAAGSLMDVL 132
CC      YQ 108 R-----AHGTEV-PCVDDERASQLYQDFFRKAKQKQLPQITVKVSASLD 149
CC      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
CC      133 HHPGMNHRVEITEGILADECAALLSDFFMRMRQEI-KAQKKAQSSTD 178
```

Search completed: November 14, 2000, 10:10:27  
Job time: 176 sec

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	530	29.6	348	2	Q9X2E8	Q9x2e8 thermotoga
2	422	23.6	357	2	Q9RY01	Q9ry01 deinococcus
3	416	23.3	241	2	Q86961	Q86961 thermotoga
4	408.5	22.8	375	2	Q84735	Q84735 chlamydia t
5	396	22.1	245	2	Q51825	Q51825 shigella fl
6	390.5	21.8	363	2	Q9REF6	Q9ref6 bartonella
7	320	17.9	599	10	Q9SYV4	Q9syv4 arabidopsis
8	300	16.8	363	10	Q9SUB7	Q9sub7 arabidopsis
9	180	10.1	224	2	Q86747	Q86747 aquifex ae
10	160	9.9	227	2	Q9REG2	Q9reg2 bartonella
11	142	7.9	345	2	Q9X3V3	Q9x3v3 pseudomonas
12	140	7.8	144	2	Q26035	Q26035 helicobacte
13	140	7.8	344	2	Q9ZJB5	Q9zjb5 helicobacte
14	139	7.8	164	2	P73717	P73717 synecocyst
15	134.5	7.5	145	2	Q9RR66	Q9rr66 deinococcus
16	127.5	7.1	376	2	Q9RKM1	Q9rkm1 streptomyce
17	125	7.0	1307	10	Q9SY10	Q9sy10 arabidopsis
18	116.5	6.5	405	3	P87241	P87241 schizosacch
19	116	6.5	201	2	Q9WYN5	Q9wyn5 thermotoga

Db 6 MKRAIELAKKGLG

Db 6 MKRAIELAKKGLGRVNPVVGAVVVKDGRITAEFGHPYFGGPHAERMAIESAR



```
QY 59 -GATIIITLEPCSHFGSTPPCVNKKIIDCKIAKVYATKD-NSLDTHGDETLRAHGIEV-E 115
Db 66 RGATLIVTLEPCDHGKTPTCTDLIIIESGIKTVIGITRDPNPVSGNGVEKFRNHGIEVIE 125
QY 116 CVDDERASQLYQDFFRAKAKQLPQITVKVYSASLDGKQANDNGOSQWITNKEVKQDVYKLR 175
Db 126 GVLEEEVKKLCEFFITVTKRPFVALKVASTLDGKIADHRGDSKWITDK-LRFKVHEMR 184
QY 176 HRHDAVLTRGRTEVLDDPOVYTRIOGKNPIKVLKSGNIHFNOQIYO--DESTPIWIY 233
Db 185 NIYSAVLGAGTVLKNPQTLCKLGRNPRVILDRKGVL--SGKVRVFEENARVIFV 242
QY 234 TENPNITSQTHIEIILYKSCDLTTILHNLKRGVCTLLVEAGTPTTSEFSIYIIDEFIL 293
Db 243 TESEE-AEYPPHVEKA-LSDCSVESILNLYERDIDSVLVEGGSKVFSEF-LDHADWVFG 299
QY 294 YAPKLLIG-GSGNYQYQYNDVIEIPDANQFEIVHSELLNQNKLTLR 340
Db 300 FYSTKIFGKGLDVFSGYLSL--VSVPP--KFKVNVNFEFSDSEFLVEMR 343

RESULT 2
Q9RY01 PRELIMINARY; PRT; 357 AA.
AC Q9RY01;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE RIBOFLAVIN-SPECIFIC DEAMINASE.
GN DR0153.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R1.
RX MEDLINE; 20036896.
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Mofatt K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zaleski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RA "Genome Sequence of the Radioresistant Bacterium Deinococcus
RT radiodurans R1.";
RL Science 286:1571-1577(1999).
DR EMBL; AF001878; AAF09742.1; -.
DR TIGR; DR0153; -.
DR INTERPRO; IPR002125; -.
DR INTERPRO; IPR002734; -.
DR PFAM; PF00383; dCMP_cyt_deam; 1.
DR PFAM; PF01872; RibD_C; 1.
DR PROSITE; PS00903; CYT_DCMP_DEAMINASES; 1.
SQ SEQUENCE 357 AA; 37226 MW; ADFBE65C10626007 CRC64;
```

```
Query Match 23.6%; Score 422; DB 2; Length 357;
Best Local Similarity 36.4%; Pred. No. 1e-25;
Matches 116; Conservative 45; Mismatches 128; Indels 30; Gaps 8;
```

```
QY 1 MDYAIQANWVGQOTGVNPPVGVAVVYNEG-----RIVGIGAHLRKGDKHAEVQALDMAQON 56
Db 12 MQALNEAAKGLRGTNPVPGVGVIVRDEGEIASEIVGRGFHPKAGEPHEAFALREAGER 71
QY 57 AEGATIIITLEPCSHFGSTPPCVNKKIIDCKIAKVYATKDNS--LDTHGDETLRAHGIEV 114
Db 72 ARGATAYVLEPCSHGORTPPCADALIAAGVARVVVAAAGDPNPQVNGRGLEKLRAGIEV 131
QY 115 ECVDDERASQLYQDFFKA-KAKOLPQITVKVYSASLDGKQANDNGOSQWITNKEVKQDVYK 173
Db 132 ATGVLEAAVAVRQAGFSLVTRGRPHVIVKYAMTLDGKVAALNEGNGPVSGPEARARVMA 191
QY 174 LRHRHDAVLTRGRTEVLDDPOVYTR-IQDGKNPIKVLKSGNIHFNOQ-----IY 223
```

```
Db 192 WRNEVDVAVGARTALDNPQLNVNRGLDGRDPRAVLDFPEGHLPASARAVREGTVLVLR 251
QY 224 ODESTPIWIYTENPNITSQTHIEIILYKSCDLTTILHNLKRGVCTLLVEAGTPTTSE- 282
Db 252 EGRSTPL---ERDPRTVILHAH-----SLQGALEQLAGLGIVATLLLEGGPTLASAF 299
QY 283 FSIYIIDEFILYAPKLLIG 301
Db 300 FEAGLIDELRVFVAPKLLG 318

RESULT 3
O86961 PRELIMINARY; PRT; 241 AA.
AC O86961;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE RIBOFLAVIN-SPECIFIC DEAMINASE (FRAGMENT).
GN RIBD.
OS Thermotoga neapolitana.
OC Bacteria; Thermotogales; Thermotoga.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Z2706-MC24;
RA Zverlov V.V., Mashchenko O.V., Liebl W., Velikodvorskaya G.A.;
RL Submitted (JUL-1998) to the EMBL/Genbank/DBJ databases.
DR EMBL; AJ009832; CAA08869.1; -.
DR INTERPRO; IPR002125; -.
DR INTERPRO; IPR002734; -.
DR PFAM; PF00383; dCMP_cyt_deam; 1.
DR PFAM; PF01872; RibD_C; 1.
DR PROSITE; PS00903; CYT_DCMP_DEAMINASES; UNKNOWN_1.
FT NON_TER 241 241
SQ SEQUENCE 241 AA; 26675 MW; 2C91552C0D897259 CRC64;

Query Match 23.3%; Score 416; DB 2; Length 241;
Best Local Similarity 40.9%; Pred. No. 1.7e-25;
Matches 97; Conservative 46; Mismatches 86; Indels 8; Gaps 6;
```

```
QY 1 MDYAIQANWVGQOTGVNPPVGVAVVYNEGRIYVIGIAH-LRKGDKHAEVQALDMAQNAE- 58
Db 6 MKRAIELAKKGLGRVNPVPPVGVAVVYKEGRIISEGFHPLFLVGPHEARVAIESAKRKGED 65
QY 59 --GATIIITLEPCSHFGSTPPCVNKKIIDCKIAKVYATKD-NSLDTHGDETLRAHGIE-V 114
Db 66 LTGTTLVVTLPECDHRHGKTPPCTDLIIIESGIRKVVIGMRDPNPVSGVGEKLLKHGIEV 125
QY 115 ECVDDERASQLYQDFFKAKAKOLPQITVKVYSASLDGKQANDNGOSQWITNKEVKQDVYK 174
Db 126 EGVLEEVKLLCEFFITVYTNRPFFALYASTLDGKIADAGNSKWIT--ODLURHAHEM 184
QY 175 RHRHDAVLTRGRTEVLDDPOVYTRIOGKNPIKVLKSGNIH-FNQOIYQDESTPI 230
Db 185 RNVSVALVAGARTVLKDDPRLTCLRKGRNPRVILDRRGLSGGSGYRVFEDNARVI 241

RESULT 4
O84735 PRELIMINARY; PRT; 375 AA.
AC O84735;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE RIBOFLAVIN DEAMINASE.
GN RIBD.
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D/UW-3/CX;
```



[illegible][illegible][illegible]



```
SQ SEQUENCE 224 AA; 25256 MW; E4A21E2EA4540024 CRC64;

Query Match
Best Local Similarity 10.1%; Score 180; DB 2; Length 224;
Matches 44; Conservative 38; Mismatches 83; Indels 12; Gaps 3;

QY 138 PQITVKVSASLDGQANDNGOS----QWITNKVKQDVYKLRHRHDVLTGRRIVELDDP 193
| : : | : | | | : : : : : | : : | : | : | : |
Db 4 PYVIIVSEVSDGKLTLRYGASSKELMSLMDDEAYKYLHEIRAKVDGIMVGCEVTRTDP 63
| : : | : | | | : : : : : | : : | : | : | : |

QY 194 QYTRIODGKNPIKVLKSKSNIHFNOOIYODESTPIWIYTEN-----PNLTSNQTHI 246
| | | | | | | | | : : : : : | : : | : | : | : |
Db 64 SLTVRYAKGNPVRILPSTGNVPLDANVLNKEAPTIIATERAPKRLKIKELGAEV 123
| : : | : | : | : | : | : | : | : | : | : | : |

QY 247 EIIYLKSCDLTTLHNLKRGVGTLLVEAGPTTTFSEF-SIYVIDEFILYAPKLIKG 302
| : : | : | : | : | : | : | : | : | : | : | : |
Db 124 IVVGDELVDKLLPELYRRIKSLMVEGGASINWFEVRRRVVDDEIRLIHLPLVIVGG 180
| : : | : | : | : | : | : | : | : | : | : | : |

RESULT 10
Q9REG2 PRELIMINARY; PRT; 227 AA.
AC Q9REG2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE RIBOFLAVIN-SPECIFIC DEAMINASE (FRAGMENT).
GN RIBD.
OS Bartonella clarridgeiae.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Bartonellaceae; Bartonella.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=73;
RA Bereswill S., Hinkelmann S., Kist M., Sander A.;
RT "Molecular analysis of riboflavin synthesis genes in bartonella
RT henselae and use of the ribC gene for differentiation of bartonella
RT species by PCR.";
RL J. Clin. Microbiol. 37:3159-3166(1999).
DR EMBL: AJ236916; CAB63087.1; -.
DR INTERPRO: IPR002734; -.
DR PFAM: PF01872; RibD_C; 1.
FT NON_TER 1
SQ SEQUENCE 227 AA; 25344 MW; CFBB8DFF6A543ED5 CRC64;

Query Match
Best Local Similarity 8.9%; Score 160; DB 2; Length 227;
Matches 59; Conservative 49; Mismatches 87; Indels 38; Gaps 11;

QY 133 KAKQLPQITVKVSASLD---GQANDNGOSQWITNKVK-QDVYKLRHRHDVLTGRTV 188
| | : : | : | : | : | : | : | : | : | : | : | : |
Db 3 KKLQRCVETLMAISADNGVCKK----GQSLKISGEISHSQTHLRAQNDVIWVGICTI 58
| | : : | : | : | : | : | : | : | : | : | : | : |

QY 189 ELDDPQYTRIQ--DGKNPIKVLKSGNIHFNOOIYODEST-PIWIYTENPLNLSNQTH 245
| | : : | : | : | : | : | : | : | : | : | : | : |
Db 59 LADDPKLDCLRPLGLEMSPRIVLDKRLRPLDAKVVQTAANIPTWVIC-GTVLSKRRKK 117
| : : | : : | : : | : : | : : | : : | : : | : : | : : |

QY 246 IEI----IYLKSCDLT-----TILHNLKRGVGTLLVEAGPTTTFSEF-SIYVIDEIL 293
| : : | : : | : : | : : | : : | : : | : : | : : | : : |
Db 118 IALEQYGVTVCSVETVNNLIPPLAILRLLYQKRLNSVLLEGAGKTGEIFLNASCVDRLIC 177
| : : | : : | : : | : : | : : | : : | : : | : : | : : |

QY 294 YYAPKLIKSGNGYQFYQTNVDVIEIPD----ANQFEIVHSELLNQNLTKRKK 342
| : : | : : | : : | : : | : : | : : | : : | : : | : : |
Db 178 FYAPLIIG-----KDRIEAPHFESYLSFNKTEMRLMGNDRLYKWRKK 220
| : : | : : | : : | : : | : : | : : | : : | : : | : : |

RESULT 11
Q9X3V3 PRELIMINARY; PRT; 145 AA.
ID Q9X3V3
AC Q9X3V3;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
```

```
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE CUMB.
GN Pseudomonas putida GB-1.
OS Bacteria; Proteobacteria; gamma subdivision; Pseudomonas group;
OC Pseudomonas.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GB-1;
RX MEDLINE; 99203138.
RA Brouwers G.J., Vrind J.P.M., Corstjens P.L.A.M., Cornelis P.,
RA Baysse C., Vrind-de Jong E.W.;
RT "cuma, a gene encoding a multicopper oxidase, is involved in Mn2+
RT oxidation in pseudomonas putida GB-1.";
RL Appl. Environ. Microbiol. 65:1762-1768(1999).
DR EMBL: AF086638; AAD24212.1; -.
DR INTERPRO: IPR002125; -.
DR PFAM: PF00383; dCMP_cyt_deam; 1.
DR PROSITE: PS00903; CYT_DCMP_DEAMINASES; UNKNOWN_1.
DR SEQUENCE 145 AA; 15171 MW; A92F44D242DFA70C CRC64;

Query Match
Best Local Similarity 7.9%; Score 142; DB 2; Length 145;
Matches 45; Conservative 24; Mismatches 54; Indels 22; Gaps 6;

QY 5 IQLANMVGQGTGVPVPGAVVNEGRIVGIGAHLRKGDGK----HAEOALDMAOQNAE-- 58
| | : : | : | : | : | : | : | : | : | : | : | : |
Db 3 LALALAAEGAALGEVPGAVLVQGVQGVGQGNRPIDSDPSAHAEMVAIRAAKASNY 62
| : : | : : | : : | : : | : : | : : | : : | : : | : : |

QY 59 ---GATYITILEPCSHFGSTPPCVNKKIIDCKIAKVYVAT---KNSLDTHGD---ETLRA 109
| : : | : : | : : | : : | : : | : : | : : | : : | : : |
Db 63 RLPGSTLYVTLEPCS-----MCAGLIVHSRVMRVFEGALEPKAGIVSQGQFGQLN 116
| : : | : : | : : | : : | : : | : : | : : | : : | : : |

QY 110 HGIEVE-CVDDERASQLYQDFFKAK 133
| : : | : : | : : | : : | : : | : : | : : | : : | : : |
Db 117 HRVIEGGVLAECACQILSDFFKAR 141
| : : | : : | : : | : : | : : | : : | : : | : : | : : |

RESULT 12
O26035 PRELIMINARY; PRT; 344 AA.
ID O26035
AC O26035;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE RIBOFLAVIN BIOSYNTHESIS PROTEIN (RIBG).
GN HPI505.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=26695 / ATCC 700392;
RX MEDLINE; 97394467.
RA Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;
RT "The complete genome sequence of the gastric pathogen Helicobacter
RT pylori.";
RL Nature 388:539-547(1997).
DR EMBL: AE000648; AAD08541.1; -.
DR TIGR: HPI505; -.
DR INTERPRO: IPR002125; -.
DR PFAM: PF00383; dCMP_cyt_deam; 1.
KW Hypothetical protein.
```





**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 14, 2000, 10:07:20 ; Search time 28.37 Seconds  
(without alignments)  
119.323 Million cell updates/sec

Title: US-08-978-456-4

Perfect score: 498

Sequence: 1 MDYATQLPNWVGXTGVNPP.....NKIIDCKTAXVVLXNXRQFR 99

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_36.\*  
1: /SIDS1/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /SIDS1/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /SIDS1/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /SIDS1/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /SIDS1/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /SIDS1/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /SIDS1/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /SIDS1/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /SIDS1/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /SIDS1/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /SIDS1/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /SIDS1/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /SIDS1/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /SIDS1/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /SIDS1/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /SIDS1/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /SIDS1/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /SIDS1/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /SIDS1/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /SIDS1/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /SIDS1/gcgdata/geneseq/geneseq/AA2000.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	488	98-0	99	18 W28084	Amino acid sequenc
2	488	98-0	99	20 Y15914	Pyrimidine deamina
3	451	90-6	342	20 Y15913	Pyrimidine deamina
4	261	52.4	361	12 R10070	B.subtilis rib ORF
5	261	52.4	1855	20 Y21803	B. subtilis rib op
6	261	52.4	1855	21 Y83271	Polypeptide encode
7	215	43.2	366	20 Y17949	S. pneumoniae ribc
8	203	40.8	401	21 Y96284	Sinorhizobium meli
9	195	39.2	170	20 Y37428	Amino acid sequenc
10	193	38.8	376	20 Y35519	Chlamydia pneumoni
11	180.5	36.2	410	19 W55031	Actinobacillus ple
12	180.5	36.2	410	20 Y21972	APP ribg gene prod

13	102.5	20.6	155	19 Y86119	S. pneumoniae deri
14	97	19.5	609	16 R88279	Ashbya gossypii DR
15	83.5	16.8	178	18 W18205	Human deoxycytidyl
16	83	16.7	344	19 W98756	H. pylori GHPO 107
17	78	15.7	591	15 R52825	DRAP-deaminase. S
18	76	15.3	158	11 R08285	Thermally stable c
19	76	15.3	158	16 R66507	yeast cytosine-dea
20	76	15.3	158	21 Y55104	Chimeric cytosine
21	76	15.3	373	20 Y53913	Amino acid sequenc
22	75.5	15.2	163	20 Y35657	Chlamydia pneumoni
23	73	14.7	158	21 Y55101	Chimeric cytosine
24	73	14.7	158	21 Y55103	Chimeric cytosine
25	69.5	14.0	176	20 Y37278	Amino acid sequenc
26	68	13.7	158	21 Y55102	Chimeric cytosine
27	66	13.3	569	16 R67378	H. pylori urease u
28	62.5	12.6	155	19 R38699	S. pneumoniae come
29	62	12.4	569	11 R04580	Part of protein w
30	62	12.4	569	17 W07194	H. pylori urease B
31	62	12.4	569	19 W98511	H. pylori GHPO 124
32	61	12.2	559	12 R13550	B subunit of H. py
33	61	12.2	749	19 W80599	Helicobacter pylor
34	61	12.2	2441	16 R79054	CREB binding prote
35	61	12.2	2441	19 W40058	Cellular transcrip
36	61	12.2	2441	21 Y94252	Mouse nuclear CREB
37	60.5	12.1	169	19 W37847	Human endokine-alp
38	60.5	12.1	169	21 Y53061	Human endokine alp
39	60.5	12.1	177	20 Y06646	Human PRO364 ligan
40	60.5	12.1	177	20 Y15817	Amino acid sequenc
41	60	12.0	346	21 Y68647	Amino acid sequenc
42	58.5	11.7	405	20 Y05469	Human CBP protein
43	58	11.6	145	19 W60119	Mycobacterium vacc
44	58	11.6	145	20 Y14865	M. vaccae antigen
45	58	11.6	273	19 W60128	M. vaccae antigen

#### ALIGNMENTS

RESULT: 1

ID W28084 standard; Protein; 99 AA.

XX W28084;

AC

XX

DT 01-SEP-1998 (first entry)

XX

DE Amino acid sequence of riboflavin biosynthesis protein RIBG.

XX

KW Staphylococcus aureus protein; ribozyme; antisense sequence; control;

KW Staphylococcal gene; regulatory element; bacterial gene expression;

KW vaccine; Staphylococcal infection; food poisoning; scaled skin syndrome;

KW toxic shock syndrome.

XX

OS Staphylococcus aureus.

XX

FH Key Location/Qualifiers

FT Misc-difference 1..99

FT /note= "residues designated X are not defined in

FT the specification"

XX

PN W09730070-AL.

XX

PD 21-AUG-1997.

XX

PF 19-FEB-1997; 97WO-US02318.

XX

PR 20-FEB-1996; 96US-0011888.

XX

PA (SMIK ) SMITHKLINE BEECHAM CORP.

XX

PI Black MT, Burham MK, Hodgson JE, Knowles DJC,

PI Pratt JM, Reichard RW, Rosenberg M, Ward JM;

XX Nicholas RO;

DR	WPI: 1997-424969/39.	DR	WPI: 1999-347572/29.
DR	N-PSDB; T84023.	DR	N-PSDB; X59910.
XX		XX	
PT	Novel polypeptide(s) from Staphylococcus aureus strain WCUH29 - used	XX	New Staphylococcus aureus polypeptide and polynucleotide useful in
PT	to isolate antimicrobial compounds, and in vaccines against S.	PT	the treatment of gastric ulcer and gastritis
PT	aureus infection	XX	
XX		XX	
PS	Claim 6; Page 454-455; 989pp; English.	PS	Claim 23; Page 6; 48pp; English.
XX		XX	
CC	The present sequence represents a Staphylococcus aureus protein, that,	CC	The present sequence is encoded by the ORF of Staphylococcus aureus
CC	based on homology with a Bacillus subtilis protein, is believed to be a	CC	pyrimidine deaminase and pyrimidine reductase (ribG) DNA. Staphylococcus
CC	riboflavin biosynthesis protein RibG. The DNA sequence was isolated from	CC	aureus ribG and its antagonists are used to treat individuals in
CC	a library of clones of S. aureus WCUH 29 in Escherichia coli. The DNA	CC	need of them. Disease related to expression or activity of ribG can
CC	sequence can be used in the construction of ribozymes and antisense	CC	be determined by analysing the nucleic acid sequence encoding ribG
CC	sequences to control the expression of Staphylococcal genes. The DNA	CC	or detecting the ribG polypeptide in a sample. ribG can also be used
CC	sequence is also useful as a source of regulatory elements for the	CC	to identify antagonists or agonists. ribG, or its related nucleic acid,
CC	control of bacterial gene expression. The present protein may be used	CC	also has use as a vaccine to induce an immunological response in an
CC	to produce vaccines to enable a host to produce specific antibodies	CC	animal. Antimicrobial compounds (e.g. agonists and antagonists of ribG),
CC	with antibacterial action. These vaccines and antibodies would protect	CC	especially broad-spectrum antibiotics, may be of use in the treatment
CC	a host against invasion by S. aureus, and conditions relating to	CC	of Helicobacter pylori infection. This should decrease the advent of
CC	Staphylococcal infection, e.g. Staphylococcal food poisoning, scaled	CC	H. pylori-induced cancers, such as gastrointestinal carcinoma. The
CC	skin syndrome, and toxic shock syndrome.	CC	treatment should also cure gastric ulcers and gastritis.
XX		XX	
SQ	Sequence 99 AA;	SQ	Sequence 99 AA;
Query Match 98.0%; Score 488; DB 18; Length 99;			
Best Local Similarity 100.0%; Pred. No. 5.5e-55;			
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1 MDYAIQLPNMVQGXGTGVPVGVVNEGRIVGIGAHLRKGDHAEVQALDMAQXNAEGA 60	QY	1 MDYAIQLPNMVQGXGTGVPVGVVNEGRIVGIGAHLRKGDHAEVQALDMAQXNAEGA 60
DB	1 mdyaiqlpnmvqgxtgvpvavvnegrivgigahlrkdgkhaevqaldmaqnaega 60	DB	1 mdyaiqlpnmvqgxtgvpvavvnegrivgigahlrkdgkhaevqaldmaqnaega 60
QY	61 TIYITLPCSHFGSTPPCVNKKIIDCKIAAXVVLXNXRQFR 99	QY	61 TIYITLPCSHFGSTPPCVNKKIIDCKIAAXVVLXNXRQFR 99
DB	61 tiyitlepcshfgstppcvnkkiiidckiaxvvlxnxrqfr 99	DB	61 tiyitlepcshfgstppcvnkkiiidckiaxvvlxnxrqfr 99
RESULT 2			
Y15914		Y15913	
ID	Y15914 standard; Protein; 99 AA.	ID	Y15913 standard; Protein; 342 AA.
XX		XX	
AC	Y15914;	AC	Y15913;
XX		XX	
DT	04-AUG-1999 (first entry)	DT	04-AUG-1999 (first entry)
XX		XX	
DE	Pyrimidine deaminase and pyrimidine reductase (ribG) encoded by ORF.	DE	Pyrimidine deaminase and pyrimidine reductase (ribG) polypeptide.
XX		XX	
KW	Pyrimidine deaminase; pyrimidine reductase; ribG; antagonist;	KW	Pyrimidine deaminase; pyrimidine reductase; ribG; antagonist;
KW	agonist; antimicrobial; antibiotic; Helicobacter pylori infection;	KW	agonist; antimicrobial; antibiotic; Helicobacter pylori infection;
KW	H. pylori-induced cancer; gastrointestinal carcinoma; gastric ulcer;	KW	H. pylori-induced cancer; gastrointestinal carcinoma; gastric ulcer;
KW	gastritis.	KW	gastritis.
XX		XX	
OS	Staphylococcus aureus.	OS	Staphylococcus aureus.
XX		XX	
PH	Key Location/Qualifiers	PH	WO9926475-A1.
FT	Misc-difference 1..99	FT	03-JUN-1999.
FT	/note= "X= any amino acid or stop codon"	XX	
XX		XX	
PN	WO9926475-A1.	PN	23-NOV-1998; 98WO-US24857.
XX		XX	
PD	03-JUN-1999.	XX	25-NOV-1997; 97US-0978456.
XX		XX	
PF	23-NOV-1998; 98WO-US24857.	XX	(SMIK ) SMITHKLINE BEECHAM CORP.
XX		XX	
PR	25-NOV-1997; 97US-0978456.	XX	Black MT, Burnham MKR, Fedon JC, Hodgson JE, Knowles DJC;
XX		XX	Lonetto MA, Nicholas RO, Palmer LM, Pratt JM, Reichard RW;
PA	(SMIK ) SMITHKLINE BEECHAM CORP.	PI	Rosenberg M, Trani CM, Ward JM, Warren RL;
PI	Black MT, Burnham MKR, Fedon JC, Hodgson JE, Knowles DJC;	PI	WPI: 1999-347572/29.
PI	Lonetto MA, Nicholas RO, Palmer LM, Pratt JM, Reichard RW;	DR	N-PSDB; X59909.
PI	Rosenberg M, Trani CM, Ward JM, Warren RL;	XX	



PT New Staphylococcus aureus polypeptide and polynucleotide useful in  
XX the treatment of gastric ulcer and gastritis  
PS Claim 8; Page 5; 48pp; English.  
XX The present sequence represents a Staphylococcus aureus pyrimidine  
CC deaminase and pyrimidine reductase (ribG) polypeptide. Staphylococcus  
CC aureus ribG and its antagonists are used to treat individuals in  
CC need of them. Disease related to expression or activity of ribG can  
CC be determined by analysing the nucleic acid sequence encoding ribG  
CC or detecting the ribG polypeptide in a sample. ribG can also be used  
CC to identify antagonists or agonists. ribG, or its related nucleic acid,  
CC also has use as a vaccine to induce an immunological response in an  
CC animal. Antimicrobial compounds (e.g. agonists and antagonists of ribG),  
CC especially broad-spectrum antibiotics, may be of use in the treatment  
CC of Helicobacter pylori infection. This should decrease the advent of  
CC H. pylori-induced cancers, such as gastrointestinal carcinoma. The  
XX treatment should also cure gastric ulcers and gastritis.  
SQ Sequence 342 AA;

Query Match 90.6%; Score 451; DB 20; Length 342;  
Best Local Similarity 95.6%; Pred. No. 1.5e-49;  
Matches 87; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 MDVAIQLPNMVQGTGVNPPGVAVVNEGRIVGIGAHLRKGDRAEVOALDMAQXNAEGA 60  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 1 mdyaiqlanmvqggtgvnppgvavvnnegrigvlgahlrkgdkhaevqaldmaagnaega 60

QY 61 TIYITLPCSHFGSTPPCVNKKIIDCKIAXV 91  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 61 tiyitlpcshfgstppcvnkkiiidckiaxv 91

RESULT 4  
R10070  
ID R10070 standard; Protein; 361 AA.  
XX  
AC R10070;  
XX  
DT 19-MAR-1991 (first entry)  
DE B.subtilis rib ORF 5 gene product.  
XX  
KW Riboflavin; vitamin B2.  
XX  
OS Bacillus subtilis.  
XX  
PN EP405370-A.  
XX  
PD 02-JAN-1991.  
XX  
PF 22-JUN-1990; 90BP-0111916.  
XX  
PR 22-JUN-1989; 89US-0370378.  
XX  
XX (HOFF ) HOFFMANN-LA ROCHE AG.  
XX  
PI Perkins JB, Pero JG, Sloma A;  
XX  
XX WPI; 1991-008756/02.  
DR Q-PSDB; Q10165.  
XX  
PT DNA encoding riboflavin biosynthetic proteins - used to produce  
PT recombinant bacteria for prodn. of riboflavin  
XX  
PS Disclosure; Fig 3; 72pp; English.  
XX  
XX The peptide is a product of the B.subtilin rib operon, involved in  
CC the biosynthesis of riboflavin.  
CC The product may be expressed in a transformed bacterial expression  
CC system for the efficient production of riboflavin (vitamin B2).

XX SQ Sequence 361 AA;

Query Match 52.4%; Score 261; DB 12; Length 361;  
Best Local Similarity 56.7%; Pred. No. 2.9e-25;  
Matches 51; Conservative 12; Mismatches 27; Indels 0; Gaps 0;

QY 1 MDVAIQLPNMVQGTGVNPPGVAVVNEGRIVGIGAHLRKGDRAEVOALDMAQXNAEGA 60  
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |  
Db 6 mklaldlakgqegqtesnplvgavvkdqqlvgmgahkygeahaevhalmagahega 65

QY 61 TIYITLPCSHFGSTPPCVNKKIIDCKIAXV 90  
||:|||||||:||||| ||: | |  
Db 66 diyvtlpcshygtppcaeliinsgikrv 95

RESULT 5  
Y21803  
ID Y21803 standard; Protein; 1855 AA.  
XX  
AC Y21803;  
XX  
DT 10-SEP-1999 (first entry)  
XX  
DE B. subtilis rib operon protein translated from reading frame 3.  
XX  
KW Riboflavin; open reading frame; ORF; structural gene; promoter;  
KW vitamin B2; Bacillus subtilis; rib operon.  
XX  
OS Bacillus subtilis.  
XX  
XX Key Location/Qualifiers  
FT Misc-difference 1..1855 /note= "residues Xaa are encoded by stop codons"  
FT  
XX US5925538-A.  
XX  
XX 20-JUL-1999.  
XX  
XX 24-AUG-1998; 98US-0138775.  
XX  
XX 11-SEP-1990; 90US-0581048.  
PR 22-JUN-1989; 89US-0370378.  
PR 21-APR-1992; 92US-0873572.  
PR 06-FEB-1995; 95US-0384626.  
PR 24-AUG-1998; 98US-0138775.  
XX  
XX (HOFF ) ROCHE VITAMINS INC.  
XX  
XX Erdenberger T, Hatch RT, Hermann T, Perkins JB;  
PI Pero JG, Sloma A;  
XX  
XX WPI; 1999-418271/35.  
DR N-PSDB; X81947.  
XX  
XX Use of bacterial strains for the over production of riboflavin  
XX  
XX Example 2; Fig 3A-S; 61pp; English.  
XX  
XX The invention relates to a method of producing riboflavin that comprises,  
CC providing a recombinant bacterium containing an exogenously introduced  
CC bacterial ORF (open reading frame) 3 (or ORF5) structural gene and an  
CC exogenous promoter, culturing the bacterium, and recovery of the product.  
CC The method is useful for the production of large quantities of riboflavin  
CC (vitamin B2). Sequences Y21801-803 are protein sequences translated from  
CC three different reading frames of B. subtilis rib operon. Y21804-806 are  
CC protein sequences translated from three different reading frames of the  
CC complementary sequence.  
XX  
XX Sequence 1855 AA;



PI Perkins JB, Pero JG, Sloma A;  
XX WPI; 2000-320439/28.  
DR N-PSDB; Z93752.  
XX  
XX  
PT Novel vector, used for the large scale production of riboflavin,  
PT comprises a bacterial or yeast riboflavin biosynthetic protein under  
PT the control of a foreign transcription element  
XX  
XX Disclosure; Figure 3; 66pp; English.  
XX  
XX Vectors comprising a nucleic acid sequence of bacterial or  
XX yeast origin, coding for one or more riboflavin biosynthetic proteins,  
XX and one or more transcription elements not naturally associated with  
XX the nucleic acid sequence are new. The vector and bacteria transformed  
XX with it are used for the large scale production (over 10g/l) of  
XX riboflavin. The riboflavin produced can be used to treat disorders  
XX associated with riboflavin deficiency such as hair loss, skin  
XX inflammation, vision deterioration and growth failure. This  
XX polypeptide is a translation of one of the three reading frames of  
XX the rib operon sequence described in GENESEQ record Z93752. See  
XX also Y83269-74.  
XX  
SQ Sequence 1855 AA;

Query Match 52.4%; Score 261; DB 21; Length 1855;  
Best Local Similarity 56.7%; Pred. No. 2.4e-24;  
Matches 51; Conservative 12; Mismatches 27; Indels 0; Gaps 0;  
QY 1 MDYAIQLPNMVQGTGVNPPVGVVNEGRIVGTGAHLRKGDKHAEVQALDMAOXNAEGA 60  
Db 372 mklaldakqegqtesnplvgavvkdgqivgmghlkygeahaevhahmagahaega 431  
QY 61 TIYITLPCSHFGSTPPCVNKKIIDCKIAXV 90  
Db 432 diyvtlpcshygtktpcaeliinsgikrv 461

RESULT 7  
Y17949  
ID Y17949 standard; Protein; 366 AA.  
AC Y17949;  
XX  
XX 04-AUG-1999 (first entry)  
XX  
XX S. pneumoniae ribG polypeptide.  
XX  
XX Streptococcus pneumoniae; pyrimidine deaminase; pyrimidine reductase;  
KW ribG; immune response; antimicrobial; H. pylori infection; cancer;  
KW gastrointestinal carcinoma; gastric ulcer; gastritis.  
XX  
OS Streptococcus pneumoniae.  
XX  
XX W09927126-A1.  
XX  
XX 03-JUN-1999.  
XX  
XX 23-NOV-1998; 98WO-US25010.  
XX  
XX 25-NOV-1997; 97US-0979616.  
XX  
XX (SMIK ) SMITHKLINE BEECHAM CORP.  
XX  
XX Black MT, Debouck C, Fedon JC, Hodgson JE, Jaworski DD;  
PI Knowles DJC, Kosmatka AL, Lonetto MA, Mooney J, Nicholas RO;  
PI Palmer LM, Shilling LK, Stodola RK, Wang M, Warren RL;  
PI Zhong Y;  
XX  
XX WPI; 1999-347726/29.  
DR N-PSDB; X77179.  
XX

PT New Streptococcus pneumoniae ribG polypeptide and polynucleotide  
XX useful in the treatment of gastric ulcer and gastritis  
XX  
XX Claim 14; Page 6; 48pp; English.  
XX  
XX The invention relates to Streptococcus pneumoniae pyrimidine deaminase  
XX and pyrimidine reductase (ribG). S. pneumoniae ribG and its antagonists a  
XX re used to treat individuals in need of the polypeptide. Disease related  
XX to expression or activity of ribG can be determined by analysing the  
XX nucleic acid sequence encoding ribG or detecting the ribG polypeptide in  
XX a sample. ribG can also be used to identify antagonists or agonists.  
XX ribG, or its related nucleic acid, also has use as a vaccine to induce  
XX an immunological response in an animal. Antimicrobial compounds (e.g.  
XX agonists and antagonists of ribG), especially broad-spectrum antibiotics,  
XX may be of use in the treatment of Helicobacter pylori infection. This  
XX should decrease the advent of H. pylori-induced cancers, such as  
XX gastrointestinal carcinoma. The treatment should also cure gastric ulcers  
XX and gastritis. The present sequence represents the S. pneumoniae ribG  
XX polypeptide.  
XX  
SQ Sequence 366 AA;

Query Match 43.2%; Score 215; DB 20; Length 366;  
Best Local Similarity 46.7%; Pred. No. 2.2e-19;  
Matches 43; Conservative 12; Mismatches 37; Indels 0; Gaps 0;  
QY 1 MDYAIQLPNMVQGTGVNPPVGVVNEGRIVGTGAHLRKGDKHAEVQALDMAOXNAEGA 60  
Db 7 mklalktaqgagvynpvmvgalivkdhlihggyheffggphaernalknckrspvga 66  
QY 61 TIYITLPCSHFGSTPPCVNKKIIDCKIAXVVL 92  
Db 67 tlyvtlpcshfghtppcdaidsgrtrvi 98

RESULT 8  
Y96284  
ID Y96284 standard; Protein; 401 AA.  
XX  
XX Y96284;  
XX  
XX 11-AUG-2000 (first entry)  
XX  
XX Sinorhizobium meliloti RibD.  
XX  
XX Plant growth; photosynthesis; alfalfa; riboflavin synthase;  
KW riboflavin synthesis; RibD.  
XX  
XX Sinorhizobium meliloti.  
XX  
XX W0200029607-A1.  
XX  
XX 25-MAY-2000.  
XX  
XX 17-NOV-1999; 99WO-US27318.  
XX  
XX 17-NOV-1998; 98US-0193600.  
XX  
XX 17-NOV-1998; 98US-0193801.  
XX  
XX 06-AUG-1999; 99US-0369955.  
XX  
XX (REGC ) UNIV CALIFORNIA.  
XX  
XX Phillips DA, Joseph CM, Sanborn JR, Yang G;  
XX  
XX WPI; 2000-387815/33.  
XX  
XX N-PSDB; A27366.  
XX  
XX Enhancing plant growth by applying an agent comprising lumichrome or  
PT lumichrome-releasing microorganism or growing plants in hydroponic  
PT culture system or in a medium comprising lumichrome-releasing  
PT microorganism -  
XX

PS Claim 61; Page 80-81; 86pp; English.

XX The present sequence is the *Sinorhizobium meliloti* RibD protein,  
CC which is involved in riboflavin synthesis. Riboflavin is important  
CC in plant growth and photosynthesis, and the protein and gene encoding it  
CC can be used to create genetically engineered microorganisms which produce  
CC an increased amount of the compound, which in turn increases plant  
CC growth. In addition to being applied to the growing plant, the bacteria  
CC can also be applied to the seed before planting. The bacteria can be  
CC applied to any plant and any part of the plant.

XX Sequence 401 AA;

Query Match 40.8%; Score 203; DB 21; Length 401;

Best Local Similarity 45.7%; Pred. No. 8.4e-18;

Matches 42; Conservative 10; Mismatches 40; Indels 0; Gaps 0;

QY 1 MDYAIQLPNVQGXGTGNPPVGAVVYVNEGRTVIGAHLRKGDHAEVQALDMAQXNAEGA 60

Db 12 maaalrlarrnlgltstnpsvgcivnkgitvgravtappgrphaetqalaegekarga 71

QY 61 TIVITLPCSHFGSTPPCVNKKIIDCKIAWVL 92

Db 72 tayvalepcshhgktppcadaliagvgrvv 103

RESULT 9

Y37428

ID Y37428 standard; Protein; 170 AA.

XX AC Y37428;

XX DT

XX DT

XX 07-OCT-1999 (first entry)

DE Amino acid sequence of a Chlamydia trachomatis protein.

XX Vaccine; eye disease; conventional trachoma; nonendemic trachoma;  
KW paratrachoma; inclusion conjunctivitis; genital disease; perihhepatitis;  
KW nongonococcal urethritis; epididymitis; cervicitis; salpingitis;  
KW bartholinitis; pneumopathy; venereal lymphogranulomatosis.

XX Chlamydia trachomatis.

OS WO928475-A2.

PN 10-JUN-1999.

XX PD

XX XX

PF 27-NOV-1998; 98WO-IB01939.

XX 04-NOV-1998; 98US-0107077.

PR 28-NOV-1997; 97FR-0015041.

PR 17-DEC-1997; 97FR-0016034.

XX (GEST ) GENSET.

PA Griffais R;

XX WPI; 1999-371125/31.

XX Genome sequence of Chlamydia trachomatis

PT Disclosure; Page 1127-1128; 1755pp; English.

XX Y36754-Y37949 are encoded by open reading frames (ORFs) of the genome of

CC Chlamydia trachomatis (see 201425). The polypeptides can be used as

CC vaccines against Chlamydia trachomatis. Antisense and ribozyme sequences

CC can also be used to control growth of the microorganism. Chlamydia

CC trachomatis is responsible for a large number of diseases, e.g. eye

CC diseases such as conventional trachoma, nonendemic trachoma,

CC paratrachoma, and inclusion conjunctivitis; genital diseases such as

CC nongonococcal urethritis, epididymitis, cervicitis, salpingitis,

CC perihhepatitis, bartholinitis; pneumopathy in breast feeding infants;

CC and venereal lymphogranulomatosis. The polypeptides of the invention  
CC may be of use in treating these diseases.

XX Sequence 170 AA;

Query Match 39.2%; Score 195; DB 20; Length 170;

Best Local Similarity 50.7%; Pred. No. 2.9e-17;

Matches 37; Conservative 11; Mismatches 25; Indels 0; Gaps 0;

QY 18 NPPVGAVVYVNEGRTVIGAHLRKGDHAEVQALDMAQXNAEGATIVITLPCSHFGSTPP 77

Db 69 npwvgcivlkgvcigegwhgigsphaevcavqdkcslegaevyvtilepcchfgrtpp 128

QY 78 CVNKKIIDCKIAWV 90

Db 129 cvdlliikskaav 141

RESULT 10

Y35519

ID Y35519 standard; Protein; 376 AA.

XX AC Y35519;

XX DT

XX 13-SEP-1999 (first entry)

XX Chlamydia pneumoniae transmembrane protein sequence.

XX Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;

XX sinusitis; purulent otitis media; erythema nodosum; pharyngitis;

XX vaccine; neutralising epitope.

XX Chlamydia pneumoniae.

XX WO927105-A2.

PN 03-JUN-1999.

XX 20-NOV-1998; 98WO-IB01890.

XX 04-NOV-1998; 98US-0107078.

PR 21-NOV-1997; 97FR-0014673.

XX (GEST ) GENSET.

XX Griffais R;

XX WPI; 1999-357842/30.

XX Genome sequence of Chlamydia pneumoniae

PS Page 1273-1274; Disclosure; 1912pp; English.

XX Y34584-Y35879 represent the proteins encoded by all the open reading

CC frames in the complete genome (see X91990) of Chlamydia pneumoniae.

CC C. pneumoniae causes respiratory disease such as pneumonia and

CC bronchitis and is thought to be a contributing factor in heart

CC disease, sarcoidosis, sinusitis, purulent otitis media, erythema

CC nodosum or pharyngitis. The polypeptides encoded by the open reading

CC frames of the C. pneumoniae genome (see Y34584-Y35879) can be used in

CC immunogenic compositions as vaccines. Vectors containing C. pneumoniae

CC nucleotide sequences can also be used as immunogenic compositions,

CC especially where the vector directs the expression of a neutralising

CC epitope of C. pneumoniae.

XX Sequence 376 AA;

Query Match 38.8%; Score 193; DB 20; Length 376;

Best Local Similarity 50.7%; Pred. No. 1.5e-16;

Matches 37; Conservative 9; Mismatches 27; Indels 0; Gaps 0;





DR	N-PSDB; T71696, T71702.
XX	
PT	DNA encoding human deoxycytidylate deaminase - for production of recombinant deaminase
PT	
XX	
XX	
PS	Disclosure; Column 109-110; 58pp; English.
XX	
CC	The present sequence represents human deoxycytidylate (dCMP) deaminase. The gene for dCMP deaminase contains a 5' untranslated region (including the promoter), 5 exons, 4 introns and a 3' untranslated region (including the stop signals). The gene can be used to produce recombinant dCMP deaminase, which can be used to convert dCMP to dUMP. Also, the dCMP gene can be altered (removed or mutated) to alter DNA replication in cells, which may lead to mutagenesis.
CC	
XX	
XX	
SQ	Sequence 178 AA;
	Query Match 16.8%; Score 83.5; DB 18; Length 178;
	Best Local Similarity 26.9%; Pred. No. 0.0053;
	Matches 28; Conservative 15; Mismatches 30; Indels 31; Gaps 5;
QY	18 NPPVGAYVVN-EGRIVGIG-----AHLRKGDK-----HAEVQA-IDM 52         :           :         :   :
Db	34 nsqvgacivnsenkivginyngmpgcddvlpwrrtaenkltdkypvychaeinaimnk 93         :           :         :   :
QY	53 AQXNAEGATITILEPCSHGSTPSPCVNKIIDCKIAVVLNXR 96 :   :   :         :         :     :
Db	94 nstdvkgsmyvalfpncne-----caklliqagikeiftsdK 131 :   :   :         :         :     :

Search completed: November 14, 2000, 10:07:25  
Job time: 43 sec

**THIS PAGE BLANK (USPTO)**



Qy 1 MDYAIQLPNWOGTGVNPPYGA VVWNEGRIVGICAHLRKGDKHAEEQALDMAQXNAEGA 60  
 |||||  
 Db 1 MDYAIQLPNWOGTGVNPPYGA VVWNEGRIVGICAHLRKGDKHAEEQALDMAQXNAEGA 60  
 |||||

```
QY 61 TITITLEPCSHFGSTPPCVNKKIIDCKIAHXVVLXNXRQFR 99
|||||
Db 61 TITITLEPCSHFGSTPPCVNKKIIDCKIAHXVVLXNXRQFR 99

RESULT 2
US-08-978-456-2
; Sequence 2, Application US/08978456
; Patent No. 6010881
; GENERAL INFORMATION:
; APPLICANT: Palmer, Leslie M.
; TITLE OF INVENTION: No. 6010881el ribG
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
; CITY: Philadelphia
; STATE: PA
; COUNTRY: US
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/978,456
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Dickinson, Todd Q.
; REGISTRATION NUMBER: 28,354
; REFERENCE/DOCKET NUMBER: P50444-9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-994-2252
; TELEFAX: 215-994-2222
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 342 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-978-456-2

Query Match 90.6%; Score 451; DB 3; Length 342;
Best Local Similarity 95.6%; Pred. No. 3e-53;
Matches 87; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 MDYAIOLPNMVQXTGVNPPVGVAVVNEGRIVGIGAHLRKGDHAEVQALDMAQXNAEGA 60
|||||
Db 1 MDYAIOLANMVQQTGVNPPVGVAVVNEGRIVGIGAHLRKGDHAEVQALDMAQXNAEGA 60

QY 61 TITITLEPCSHFGSTPPCVNKKIIDCKIAHV 91
|||||
Db 61 TITITLEPCSHFGSTPPCVNKKIIDCKIAKV 91

RESULT 3
US-08-741-327E-15
; Sequence 15, Application US/08741327E
; Patent No. 5925354
; GENERAL INFORMATION:
; APPLICANT: Fuller, et al.
; TITLE OF INVENTION: Riboflavin mutants as vaccines
; TITLE OF INVENTION: against Actinobacillus pleuropneumoniae
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
```

```
; ADDRESSEE: G. Kenneth Smith
; STREET: 300 S. Wacker Drive
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/741,327E
; FILING DATE: October 28, 1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: G. Kenneth Smith
; REGISTRATION NUMBER: 43,135
; REFERENCE/DOCKET NUMBER: 97704-E
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-9808
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 410 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-741-327E-15

Query Match 36.2%; Score 180.5; DB 2; Length 410;
Best Local Similarity 43.2%; Pred. No. 2.1e-16;
Matches 41; Conservative 11; Mismatches 40; Indels 3; Gaps 1;

QY 1 MDY---AIOLPNMVQXTGVNPPVGVAVVNEGRIVGIGAHLRKGDHAEVQALDMAQXNA 57
|||||
Db 50 LDYMRRAIALAKOGLGWTNPPLVGCIVKNGEIVAEVYHEKIGGWHAERNAVLHCKEDL 109

QY 58 EGATITITLEPCSHFGSTPPCVNKKIIDCKIAHV 92
|||||
Db 110 SGATAYVTLEPCCHHGTPTPCSDLIERGIKVF 144

RESULT 4
US-08-716-301-4
; Sequence 4, Application US/08716301
; Patent No. 5821090
; GENERAL INFORMATION:
; APPLICANT: Doval, Jose Luis Revuelta
; APPLICANT: Serna, Maria Jose Buitago
; APPLICANT: Gareiss, Maria Angeles
; TITLE OF INVENTION: Riboflavin biosynthesis in fungi
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Keil & Weinkauff
; STREET: 1101 Connecticut Avenue
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS version 7.0
; SOFTWARE: Wordperfect version 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/716,301
; FILING DATE: 24-SEP-1996
; CLASSIFICATION: 435
; CLASSIFICATION: C 12 N 15/53
; CLASSIFICATION: C 12 N 15/54
; CLASSIFICATION: C 12 N 15/55
```





NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US96/05800  
FILING DATE: 23-APR-1996  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/431,041  
FILING DATE: 28-APR-1995  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/568,122  
FILING DATE: 06-DEC-1995  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Clark, Paul T.  
REGISTRATION NUMBER: 30,162  
REFERENCE/DOCKET NUMBER: 06137/020001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 566 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US96-05800-3

Query Match 12.4%; Score 62; DB 4; Length 566;  
Best Local Similarity 22.1%; Pred. No. 4;  
Matches 21; Conservative 14; Mismatches 40; Indels 20; Gaps 3;  
Qy 3 YAIOLPNMVCGXTGVPNPVGVVNEGR-----IVGIGHLRKGDKHAEVQA 49  
Db 149 FASGVTTMIGGCTGPDCTNATTITPGRNLKWLRAEEYSMLG-FLAKGNASNDASL 207  
Qy 50 LDMAQXNAEGATITILEPCSHFGSTPPCVNKIID 84  
Db 208 ADQIEAGAIGFKIH-----EDWGTTPSAINHLD 236

RESULT 11  
us-08-194-468-2  
Sequence 2, Application US/08194468  
Patent No. 5750336  
GENERAL INFORMATION:  
APPLICANT: Montminy, Marc R.  
TITLE OF INVENTION: ASSAYS FOR THE IDENTIFICATION OF  
TITLE OF INVENTION: COMPOUNDS WHICH INHIBIT ACTIVATION OF CAMP AND MITOGEN  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark  
STREET: 444 South Flower Street, Suite 2000  
CITY: Los Angeles  
STATE: California  
COUNTRY: USA  
ZIP: 90071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/194,468  
FILING DATE: 10-FEB-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Reiter, Stephen E.  
REGISTRATION NUMBER: 31,192  
REFERENCE/DOCKET NUMBER: P41 9672  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619)-546-4737  
TELEFAX: (619)-546-9392  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2441 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-194-468-2

Query Match 12.2%; Score 61; DB 1; Length 2441;  
Best Local Similarity 25.3%; Pred. No. 40;  
Matches 21; Conservative 14; Mismatches 32; Indels 16; Gaps 3;  
Qy 6 OLPNMVCGXTG--VNPPVGVVNEGRIVGIGHLRKGDKHAEVQALDMAQXNAEG---- 59  
Db 68 QLSPELLRGSGSSINPGIGNVSASSPVQOGLG-----QQAQQPNSTNNASLGAMGKSPL 122  
Qy 60 -----ATIVITLPCSHFGSTPP 77  
Db 123 NQGDSTPNLPKQAASTSGTTPP 145

RESULT 12  
US-08-961-739-2  
Sequence 2, Application US/08961739A  
Patent No. 6063583  
GENERAL INFORMATION:  
APPLICANT: Montminy, Marc R.  
TITLE OF INVENTION: Methods for Treating Diabetes Mellitus  
FILE REFERENCE: SALK1650-1  
CURRENT APPLICATION NUMBER: US/08/961.739A  
CURRENT FILING DATE: 1997-10-31  
EARLIER APPLICATION NUMBER: US 194,468  
EARLIER FILING DATE: 1994-02-10  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 2  
LENGTH: 2441  
TYPE: PRT  
ORGANISM: Mus  
FEATURE:  
NAME/KEY: VARIANT  
LOCATION: (1)...(2441)  
OTHER INFORMATION: Xaa = Any Amino Acid  
US-08-961-739-2

Query Match 12.2%; Score 61; DB 3; Length 2441;  
Best Local Similarity 25.3%; Pred. No. 40;  
Matches 21; Conservative 14; Mismatches 32; Indels 16; Gaps 3;  
Qy 6 OLPNMVCGXTG--VNPPVGVVNEGRIVGIGHLRKGDKHAEVQALDMAQXNAEG---- 59  
Db 68 QLSPELLRGSGSSINPGIGNVSASSPVQOGLG-----QQAQQPNSTNNASLGAMGKSPL 122  
Qy 60 -----ATIVITLPCSHFGSTPP 77  
Db 123 NQGDSTPNLPKQAASTSGTTPP 145

```
RESULT 13
US-08-912-227-2
; Sequence 2, Application US/08912227
; Patent No. 598171
; GENERAL INFORMATION:
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Ni, Jian
; APPLICANT: Rosen, Craig A.
; TITLE OF INVENTION: Human Endokine Alpha
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C
; STREET: 1100 NEW YORK AVE., NW, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/912,227
; FILING DATE: Herewith
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/024,058
; FILING DATE: 16-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Steffe, Eric K.
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.0470001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 169 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-912-227-2

Query Match 12.1%; Score 60.5; DB 2; Length 169;
Best Local Similarity 43.8%; Pred. No. 1.2;
Matches 14; Conservative 3; Mismatches 4; Indels 11; Gaps 2;

QY 67 EPC-SHFG-----STPPCVNKIIDCKI 87
||| : || | ||||| : | :
Db 48 EPCMAKEGLPSKQWQMASSEPPCVNKVSDKWL 79

RESULT 14
US-08-997-080-55
; Sequence 55, Application US/08997080
; Patent No. 598524
; GENERAL INFORMATION:
; APPLICANT: WATSON, JAMES D.
; APPLICANT: TAN, PAUL L.J.
; TITLE OF INVENTION: METHODS AND COMPOUNDS FOR THE TREATMENT OF IMMUNOLOGICALLY-
; NUMBER OF SEQUENCES: 194
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Offices of Ann W. Speckman
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/997,362
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: U.S. Patent Application No. 5985287 08/873,970
```

```
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/997,080
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sleath, Janet
; REGISTRATION NUMBER: 37,007
; REFERENCE/DOCKET NUMBER: 11000.1007
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-269-0565
; TELEFAX: 206-269-0563
; TELEX:
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 145 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-997-080-55

Query Match 11.6%; Score 58; DB 2; Length 145;
Best Local Similarity 28.2%; Pred. No. 2.2;
Matches 22; Conservative 8; Mismatches 22; Indels 26; Gaps 4;

QY 19 PPVGAVVNVNEGRI-----VCIGAHLRK-GDKHAEVQALDMAQX 55
||| : ||| : ||||| : | : || :
Db 58 PAVDAVVVSNDDYDHLDDITVALAHTQRAFPVPLVGIAHLRKGVPPEARIVELDMHEA 117

QY 56 N-AEGATYITILEPSCSHF 72
: : | : | | |
Db 118 HRIDDLTLVCT--PARHF 133

RESULT 15
US-08-997-362-55
; Sequence 55, Application US/08997362
; Patent No. 5985287
; GENERAL INFORMATION:
; APPLICANT: Tan, Paul
; APPLICANT: Hiyama, Jun
; APPLICANT: Visser, Elizabeth
; APPLICANT: Skinner, Margot
; APPLICANT: Scott, Linda
; APPLICANT: Prestidge, Ross
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR
; NUMBER OF SEQUENCES: 194
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Offices of Ann W. Speckman
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/997,362
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: U.S. Patent Application No. 5985287 08/873,970
```

Search completed: November 14, 2000, 10:07:47  
Job time: 65 sec

**THIS PAGE BLANK (USPTO)**







A:Molecule type: DNA  
A:Residues: 1-369 <TT>  
A:Cross-references: GB:AEO02531; GB:AEO02098; NID:g7227065; PIDN:AAF42152.1; PID:g7227070  
A:Experimental source: serogroup B, strain MMD58  
C:Genetics:  
A:Gene: NMB1817  
C:Superfamily: Chlamydochilla pneumoniae riboflavin-specific deaminase

Query Match 43.0%; Score 214; DB 2; Length 369;  
Best Local Similarity 44.0%; Pred. No. 1.2e-16;  
Matches 40; Conservative 19; Mismatches 32; Indels 0; Gaps 0;

QY 1 MDYAIIQLPNVVGXTGYNPPGVAVVNEGRIVTGAIHLRKGDKHAEVQALDMAOXNAEGA 60  
I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :  
Db 10 MENALRLAALGRFSTSPNPRVCGCIANGSQIVGGCFHVKAGEPHAEVHALRQAEGMAQGA 69  
I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :  
QY 61 TIYTILEPCSHFGSTPPCVNKIIDCKTAXVV 91  
I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :  
Db 70 TAFVTLEPCSHYGRTPPCAENALVRAGVSRRV 100  
I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :

RESULT 6  
H64103  
Riboflavin-specific deaminase (EC 3.5.4.-) - Haemophilus influenzae (strain Rd KW20)  
C:Species: Haemophilus influenzae  
C>Date: 18-Aug-1995 #sequence\_revision 18-Aug-1995 #text\_change 20-Apr-2000  
C:Accession: H64103  
R:Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodok, A.; Kelley, J.M.; Weidman, J.; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M. Science 269, 496-512, 1995  
A:Authors: Gnehm, C.U.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, A.; Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.  
A:Reference number: A64000; MUID:95350630  
A:Accession: H64103  
A>Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-372 <TIGR>  
A:Cross-references: GB:U32775; GB:L42023; NID:g1573951; PIDN:AAC22598.1; PID:g1573964; T  
C:Genetics:  
A:Gene: ribD; ribG  
C:Function:  
A:Description: converts 2,5-diamino-6-(ribosylamino)-4(3H)-pyrimidinone 5'-phosphate int  
A:Pathway: riboflavin biosynthesis  
C:Superfamily: Chlamydochilla pneumoniae riboflavin-specific deaminase  
C:Keywords: hydrolase; riboflavin biosynthesis; zinc  
F:55,80,89/Binding site: zinc (His, Cys, Cys) #status predicted

Query Match 43.0%; Score 214; DB 2; Length 372;  
Best Local Similarity 47.3%; Pred. No. 1.2e-16;  
Matches 43; Conservative 11; Mismatches 37; Indels 0; Gaps 0;

QY 1 MDYAIIQLPNVVGXTGYNPPGVAVVNEGRIVTGAIHLRKGDKHAEVQALDMAOXNAEGA 60  
I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :  
Db 12 MORALDLAAKGQYTTTPNSPGCVLVNGEIVEGSGFFKAGOPHAERVALAQACENAKGA 71  
I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :  
QY 61 TIYTILEPCSHFGSTPPCVNKIIDCKTAXVV 91  
I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :  
Db 72 TAYVTLEPCAHYGRTPPCALGLIEAGVVKVI 102  
I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :

RESULT 7  
G75552  
Riboflavin-specific deaminase - Deinococcus radiodurans (strain R1)  
C:Species: Deinococcus radiodurans  
C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 20-Apr-2000  
C:Accession: G75552  
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma S.; Smith, H.O.; Venter, J.C.; Fraser, C.M. Science 286, 1571-1577, 1999



A:Gene: ribD  
C:Superfamily: Chlamydophila pneumoniae riboflavin-specific deaminase

Query Match 25.1%; Score 125; DB 2; Length 336;  
Best Local Similarity 31.7%; Pred. No. 1.3e-06;  
Matches 33; Conservative 16; Mismatches 29; Indels 26; Gaps 3;

QY 1 MDYAIQLPNMVQSGXTGVNPPVGAVVWNEG-RIVGICAHLRKGDKHAEEVQALDMA----- 53

Db 16 MNIALNEAWKYQFLTYPNPAVGCVILDKNEKILATKAHEKAGLAHAETNAIAHAFKSLRP 75

QY 54 -----QXNA-----EGATIIITLEPCSHFGSTPPC 78  
          .     ||       . . . . .

Db 76 EISLPKEANALHEFFICKNHOGVFKDSIAFVTLEPCSHOGKTPPC 119

Search completed: November 14, 2000, 10:08:13  
Job time: 91 sec

Query Match	37.3%	Score	186;	DB	2;	Length	396;
Best Local Similarity	48.0%;	Pred. No.	1.9e-13;				
Matches	36;	Conservative	10;	Mismatches	29;	Indels	0;
						Gaps	0;

**Qy** 18 NPPGAVVNVNEGRIVIGIAHLRKGDKHAIEVQALDMAQXNAEGATIITLEPCSHFGSTPP 77  
||| | : | : | | | |||| | : : : |||| | | |  
**Db** 54 NPWGCIVYKNCGIIRGWHRHGKGISPAEVCFAFODOTSSIVAGADVVTILEPCACHEGRTPP 11

Qv 78 CYNKIIDCKIAXVWL 92

QY 78 CVNKIIDCKTAXVLT 92  
11: : 1 : : 1 :

DB 114 CVDI.I.TKSKVSSVVT 128

UD 114 CVDLLTKSKVSSVYI 128

RESULT 15

RESULT  
D81258

D81238  
probable riboflavin-specific deaminase (EC 3.5.4.-) C11622 (imported) - Cammylobacter

*C. species*: Camovlobacter tojuni

C;Species: Campylobacter jejuni  
C;Date: 31-Mar-2000 #sequences read

```
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 31-Mar-2000
C;Accession: D81258
```

C;Accession: D81258

R; Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chillingworth, C.W. *Genome* 2001, 14, 1048-1057.

C.W.; Quail, M.; Rajandre

Nature 403, 665-668, 2000

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 14, 2000, 10:10:27 ; Search time 16.17 Seconds  
(without alignments)  
195.600 Million cell updates/sec

Title: US-08-978-456-4  
Perfect score: 498  
Sequence: 1 MDYAIQLPNMVGXTGVNPP.....NKIIDCKTAXVVLXNXRQFR 99

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 87993 seqs, 31947931 residues

Total number of hits satisfying chosen parameters: 87993

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_39.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	261	52.4	361	1 RIBD_BACSU	P17618 b riboflavi
2	249	50.0	356	1 RIBD_AQUAE	O66534 a riboflavi
3	243	48.8	371	1 RIBD_BACAM	P70814 b riboflavi
4	229	46.0	367	1 RIBD_ECOLI	P25539 e riboflavi
5	216.5	43.5	339	1 RIBD_MYCTU	P71677 m riboflavi
6	214	43.0	372	1 RIBD_HAEIN	P44326 h riboflavi
7	196	39.4	368	1 RIBD_SYNY3	O55158 s riboflavi
8	195.5	39.3	337	1 RIBD_CORAM	O24750 c riboflavi
9	193	38.8	376	1 RIBD_CHLPN	Q92735 c riboflavi
10	180.5	36.2	376	1 RIBD_ACTPL	P50853 a riboflavi
11	117.5	23.6	161	1 YAAJ_BACSU	P21335 bacillus su
12	107	21.5	173	1 YFHC_HAEIN	P44931 haemophilus
13	105.5	21.2	178	1 YFHC_ECOLI	P30134 escherichia
14	96	19.3	151	1 Y903_AQUAE	O67050 aquifex aeo
15	88.5	17.8	147	1 YLXG_VIBFI	P33968 vibrio fisc
16	88	17.7	188	1 DCTD_BPT2	P00814 bacterioph
17	88	17.7	193	1 DCTD_BPT4	P16006 bacterioph
18	83.5	16.8	178	1 DCTD_HUMAN	P32321 homo sapien
19	78	15.7	591	1 RIB2_YEAST	Q12362 saccharomyc
20	76	15.3	158	1 FCY1_YEAST	Q12178 saccharomyc
21	71.5	14.4	148	1 Y831_RICPR	Q92cc6 rickettsia
22	70.5	14.2	128	1 DCTD_BPM2D	O22000 mycobacteri
23	69.5	14.0	197	1 DCTD_CAEEL	P30648 caenorhabdi
24	68	13.7	203	1 DCTD_DROME	O9vwa2 drosophila
25	66	13.3	568	1 URE2_HELHE	P42823 helicobacte
26	65.5	13.2	342	1 DCTD_YEAST	P06773 saccharomyc
27	64.5	13.0	150	1 FCAL_CANAL	P78594 candida alb
28	64.5	13.0	865	1 SECA_HELPY	O25475 helicobacte
29	63.5	12.8	211	1 HFA5_HAEIN	P45990 haemophilus
30	62.5	12.6	186	1 YD66_AQUAE	O67378 aquifex aeo
31	62	12.4	569	1 URE2_HELPY	P14917 helicobacte
32	61	12.2	352	1 RLA0_HALCU	P17006 halobacteri
33	61	12.2	352	1 RLA0_HALHA	P13553 halobacteri

RESULT\_1

ID	RIBD_BACSU	STANDARD;	PRT;	361 AA.
AC	P17618;			
DT	01-AUG-1990 (Rel. 15, Created)			
DT	01-AUG-1990 (Rel. 15, Last sequence update)			
DT	30-MAY-2000 (Rel. 39, Last annotation update)			
DE	RIBOFLAVIN BIOSYNTHESIS PROTEIN RIBD {INCLUDES: DIAMINOHYDROXYPHOSPHORIBOSYLAMINOPYRIMIDINE DEAMINASE (EC 3.5.4.26) (RIBOFLAVIN-SPECIFIC DEAMINASE); 5-AMINO-6-(5- PHOSPHORIBOSYLAMINO)URACIL REDUCTASE (EC 1.1.1.193) (HTP REDUCTASE)}.			
GN	RIBD OR RIBG.			
OS	Bacillus subtilis.			
OC	Bacteria; Firmicutes; Bacillus/Clostridium group;			
OC	Bacillus/Staphylococcus group; Bacillus.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=168 / MAREBURG;			
RX	MEDLINE; 95020538.			
RA	Sorokin A.V., Zumstein E., Azevedo V., Ehrlich S.D., Serror P.;			
RT	"The organization of the Bacillus subtilis 168 chromosome region between the spoVA and serA genetic loci, based on sequence data.";			
RL	Mol. Microbiol. 10:385-395(1993).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=168 / SHGW;			
RA	Mironov V.N.;			
RT	Thesis (1989), USSR Academy of Sciences, Russia.			
RN	[3]			
RP	SEQUENCE OF 1-12 FROM N.A.			
RC	STRAIN=168 / SHGW;			
RX	MEDLINE; 90271920.			
RA	Mironov V.N., Perumov D.A., Kraev A.S., Stepanov A.I., Skryabin K.G.;			
RT	"Unusual structure of the regulatory region of the riboflavin biosynthesis operon in Bacillus subtilis.";			
RL	Mol. Biol. (Mosk) 24:256-261(1990).			
RN	[4]			
RP	CHARACTERIZATION.			
RC	MEDLINE; 97221604.			
RA	Richter G., Fischer M., Krieger C., Eberhardt S., Luttgen H.,			
RT	Gerstenschlager I., Bacher A.;			
RT	"Biosynthesis of riboflavin: Characterization of the bifunctional deaminase-reductase of Escherichia coli and Bacillus subtilis.";			
RL	J. Bacteriol. 179:2022-2028(1997).			
CC	-1- FUNCTION: CONVERTS 2,5-DIAMINO-6-(RIBOSYLAMINO)-4(3H)-PYRIMIDINONE 5'-PHOSPHATE INTO 5-AMINO-6-(RIBOSYLAMINO)-2,4(1H,3H)- PYRIMIDINEDIONE 5'-PHOSPHATE.			
CC	-1- CATALYTIC ACTIVITY: 2,5-DIAMINO-6-HYDROXY-4-(5- PHOSPHORIBOSYLAMINO)PYRIMIDINE + H(2)O = 5-AMINO-6-(5- PHOSPHORIBOSYLAMINO)URACIL + NH(3).			
CC	-1- CATALYTIC ACTIVITY: 5-AMINO-6-(5-PHOSPHORIBITVYLAMINO)URACIL + NADP(+) = 5-AMINO-6-(5-PHOSPHORIBOSYLAMINO)URACIL + NADPH.			
CC	-1- COFACTOR: ZINC (BY SIMILARITY).			
CC	-1- PATHWAY: SECOND AN THIRD STEPS OF RIBOFLAVIN BIOSYNTHESIS.			
CC	-1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE CYTIDINE AND DEOXYCYTIDYLATE DEAMINASES FAMILY.			

Q43207 triticum ae  
P45481 mus musculu  
P47058 saccharomyc  
Q9zep8 pseudomonas  
P32393 bacillus su  
P25939 epstein-bar  
P05674 venezuelan  
P09592 venezuelan  
P36329 venezuelan  
P36332 venezuelan  
P34401 caenorhabdi  
Q9rvd3 deinococcus

ALIGNMENTS

```

CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE HTP
CC REDUCTASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L09228; AAA67481.1; -.
CC DR EMBL; X51510; CAA35878.1; -.
CC DR EMBL; Z99116; CAB14260.1; -.
CC DR PIR; A35711; A35711.
CC DR PIR; S45543; S45543.
CC DR SUBTILIST; BG10518; RIBD.
CC DR INTERPRO; IPR002125; -.
CC DR INTERPRO; IPR002734; -.
CC DR PFAM; PF01872; RIBD_C; 1.
CC DR PFAM; PF00383; dCMP_cyt_deam; 1.
CC DR PROSITE; PS00903; CYT_DCMP_DEAMINASES; 1.
CC KW Riboflavin biosynthesis; Hydrolase; Zinc; Oxidoreductase; NADP;
CC Multifunctional enzyme.
CC FT DOMAIN 1 144 DEAMINASE.
CC FT METAL 145 361 REDUCTASE.
CC FT METAL 49 49 ZINC (BY SIMILARITY).
CC FT METAL 74 74 ZINC (BY SIMILARITY).
CC FT METAL 83 83 ZINC (BY SIMILARITY).
CC SQ SEQUENCE 361 AA; 39305 MW; DA836930BFDECA3B CRC64;

Query Match 52.4%; Score 261; DB 1; Length 361;
Best Local Similarity 56.7%; Pred. No. 3.8e-22;
Matches 51; Conservative 12; Mismatches 27; Indels 0; Gaps 0;

QY 1 MDYAIOLPNVQGTGVNPPVGVAVVNEGRVIGIGHLRKGRKHAQVQALDQAQXNAEGA 60
   I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
Db 6 MKLALDLAKOGEGQTESNPLVGVAVVKGDIQVCGAHLKYGAEHAEVHAHMAHAEGA 65

QY 61 TIYITLPCSHFGSTPPCVNKKIIDCKIAXV 90
   I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
Db 66 DIYVTLPCSHYGTTPPCAEIILNSGIKRV 95

RESULT 2
RIBD_AQUAE STANDARD; PRT; 356 AA.
AC 066534;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE RIBOFLAVIN BIOSYNTHESIS PROTEIN RIBD [INCLUDES:
DE DIAMINOHYDROXYPHOSPHORIBOSYLAMINOPYRIMIDINE DEAMINASE (EC 3.5.4.26)
DE (RIBOFLAVIN-SPECIFIC DEAMINASE); 5-AMINO-6-(5-
DE PHOSPHORIBOSYLAMINO)URACIL REDUCTASE (EC 1.1.1.193) (HTP REDUCTASE)].
GN RIBD OR RIBS OR AQ_138.
OS Aquifex aeolicus.
OC Bacteria; Aquificales; Aquificaceae; Aquifex.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VF5;
RX MEDLINE; 98196666.
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Aujay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
RT aeolicus.";
RL Nature 392:353-358(1998).
CC -!- FUNCTION: CONVERTS 2,5-DIAMINO-6-(RIBOSYLAMINO)-4(3H)-PYRIMIDINONE
CC 5'-PHOSPHATE INTO 5-AMINO-6-(RIBOSYLAMINO)-2,4(1H,3H)-
CC PYRIMIDINEDIONE 5'-PHOSPHATE.
CC -!- CATALYTIC ACTIVITY: 2,5-DIAMINO-6-HYDROXY-4-(5-

```

```

CC PHOSPHORIBOSYLAMINO)PYRIMIDINE + H(2)O = 5-AMINO-6-(5-
CC CATALYTIC ACTIVITY: 5-AMINO-6-(5-PHOSPHORIBITYLAMINO)URACIL +
CC NADP(+) = 5-AMINO-6-(5-PHOSPHORIBOSYLAMINO)URACIL + NADPH.
CC -!- COFACTOR: ZINC (BY SIMILARITY).
CC -!- PATHWAY: SECOND AN THIRD STEPS OF RIBOFLAVIN BIOSYNTHESIS.
CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE CYTIDINE AND
CC DEOXYCYTIDYLATE DEAMINASES FAMILY.
CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE HTP
CC REDUCTASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE000675; AAC06487.1; -.
CC DR INTERPRO; IPR002125; -.
CC DR INTERPRO; IPR002734; -.
CC DR PFAM; PF01872; RIBD_C; 1.
CC DR PFAM; PF00383; dCMP_cyt_deam; 1.
CC DR PROSITE; PS00903; CYT_DCMP_DEAMINASES; 1.
CC KW Riboflavin biosynthesis; Hydrolase; Zinc; Oxidoreductase; NADP;
CC Multifunctional enzyme.
CC FT DOMAIN 1 148 DEAMINASE.
CC FT METAL 149 356 REDUCTASE.
CC FT METAL 53 53 ZINC (BY SIMILARITY).
CC FT METAL 78 78 ZINC (BY SIMILARITY).
CC FT METAL 87 87 ZINC (BY SIMILARITY).
CC SQ SEQUENCE 356 AA; 39628 MW; 7775515B1E4D02BD CRC64;

Query Match 50.0%; Score 249; DB 1; Length 356;
Best Local Similarity 54.3%; Pred. No. 8.3e-21;
Matches 50; Conservative 11; Mismatches 31; Indels 0; Gaps 0;

QY 1 MDYAIOLPNVQGTGVNPPVGVAVVNEGRVIGIGHLRKGRKHAQVQALDQAQXNAEGA 60
   I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
Db 10 MKLALDLAKOGEGQTESNPLVGVAVVKGDIQVCGAHLKYGAEHAEVHAHMAHAEGA 69

QY 61 TIYITLPCSHFGSTPPCVNKKIIDCKIAXV 92
   I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
Db 70 TLXVTLPCSHFGSTPPCTDAILRSIGIKRVV 101

RESULT 3
RIBD_BACAM STANDARD; PRT; 371 AA.
AC P70814;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE RIBOFLAVIN BIOSYNTHESIS PROTEIN RIBD [INCLUDES:
DE DIAMINOHYDROXYPHOSPHORIBOSYLAMINOPYRIMIDINE DEAMINASE (EC 3.5.4.26)
DE (RIBOFLAVIN-SPECIFIC DEAMINASE); 5-AMINO-6-(5-
DE PHOSPHORIBOSYLAMINO)URACIL REDUCTASE (EC 1.1.1.193) (HTP REDUCTASE)].
GN RIBD OR RIBG.
OS Bacillus amyloliquefaciens.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A 50;
RX MEDLINE; 97376523.
RA Gusarov I.I., Krenova R.A., Podcharniaev D.A., Iomantas I.U.V.,
RA Abalagina E.G., Stoinova N.V., Perumov D.A., Kozlov I.U.I.;
RT "Riboflavin biosynthetic genes in Bacillus amyloliquefaciens: primary
RT structure, organization and regulation of activity.";
RL Mol. Biol. (Mosk) 31:446-453(1997).
CC -!- FUNCTION: CONVERTS 2,5-DIAMINO-6-(RIBOSYLAMINO)-4(3H)-PYRIMIDINONE

```



mutation.";  
Mol. Gen. Genet. 234:429-432(1992).  
[2]  
SEQUENCE FROM N.A.  
STRAIN=K12 / MG1655;  
MEDLINE; 97426617.  
Blietner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
Mau B., Shao Y.;  
"The complete genome sequence of *Escherichia coli* K-12.";  
Science 277:1453-1474 (1997).  
[3]  
SEQUENCE FROM N.A.  
Roberts D., Allen E., Araujo R., Aparicio A., Chung E., Davis K.,  
Duncan M., Federspiel N., Hyman R., Kalman S., Komp C., Kurdi O.,  
Lew H., Lin D., Namath A., Oefner P., Schramm S., Davis R.W.;  
Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.  
[4]  
CHARACTERIZATION.  
MEDLINE; 97221604.  
Richter G., Fischer M., Krieger C., Eberhardt S., Luttgen H.,  
Gerstenschlager I., Bacher A.;  
"Biosynthesis of riboflavin: Characterization of the bifunctional  
deaminase-reductase of *Escherichia coli* and *Bacillus subtilis*.";  
J. Bacteriol. 179:2022-2028(1997).  
CC -1- FUNCTION: CONVERTS 2,5-DIAMINO-6-(RIBOSYLAMINO)-4(3H)-PYRIMIDINONE  
CC 5'-PHOSPHATE INTO 5-AMINO-6-(RIBOSYLAMINO)-2,4(1H,3H)-  
CC PYRIMIDINEDIONE 5'-PHOSPHATE.  
CC -1- CATALYTIC ACTIVITY: 2,5-DIAMINO-6-HYDROXY-4-(5-  
CC PHOSPHORIBOSYLAMINO)PYRIMIDINE + H(2)O = 5-AMINO-6-(5-  
CC PHOSPHORIBOSYLAMINO)URACIL + NH(3).  
CC -1- CATALYTIC ACTIVITY: 5-AMINO-6-(5-PHOSPHORIBITYLAMINO)URACIL +  
CC NADP(+) = 5-AMINO-6-(5-PHOSPHORIBOSYLAMINO)URACIL + NADPH.  
CC -1- COFACTOR: ZINC (BY SIMILARITY).  
CC -1- PATHWAY: SECOND AN THIRD STEPS OF RIBOFLAVIN BIOSYNTHESIS.  
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE CYTIDINE AND  
CC DEOXYCYTIDYLATE DEAMINASES FAMILY.  
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE HTP  
CC REDUCTASE FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial/  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; X64395; CAA45735.1; -.  
DR EMBL; AE000148; AAC73517.1; -.  
DR EMBL; U82664; AAB40170.1; -.  
DR PIR; S26201; S26201.  
DR ECOGENE; EGI1321; RIBD.  
DR INTERPRO; IPR002125; -.  
DR INTERPRO; IPR002734; -.  
DR PFAM; PF01872; RIBD\_C.1.  
DR PFAM; PF00383; dCMP\_cyt\_deam.1.  
DR PROSITE; PS00903; CYT\_dCMP\_DEAMINASES; 1.  
DR Riboflavin biosynthesis; Hydrolyase; Zinc; Oxidoreductase; NADP;  
KW Multifunctional enzyme.  
CC DOMAIN 1 145 DEAMINASE.  
FT DOMAIN 146 367 REDUCTASE.  
FT METAL 50 50 ZINC (BY SIMILARITY).  
FT METAL 75 75 ZINC (BY SIMILARITY).  
FT METAL 84 84 ZINC (BY SIMILARITY).  
SQ SEQUENCE 367 AA; 40338 MW; B19CEFA74D48D14D CRC64;

Query Match 46.0%; Score 229; DB 1; Length 367;  
Best Local Similarity 51.6%; Pred. No. 1.5e-18;  
Matches 47; Conservative 11; Mismatches 33; Indels 0; Gaps 0;

```
QY 1 MDVAIQLPNMVQGTGVNPPVGVAVVNEGRIVGIGAHLRKGDHAEVQALDMAQXNAEGA 60
  1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 7 MARALKKAQRGRTTHPNPNVNCVIVVDGELVGVGHQAGEPHAEVHALRMAKEKAKGA 66
  1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 61 TIXITLPCSHFGSTGTPPCVNKIIDCKIAXV 91
  1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 67 TAYVTLEPCSHHGRTPPCCDALTAAGVARV 97
  1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 5
RIBD_MYCTU
ID RIBD_MYCTU STANDARD; PRT; 339 AA.
AC P71677;
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE RIBOFLAVIN BIOSYNTHESIS PROTEIN RIBD [INCLUDES:
DE DIAMINOHYDROXYPHOSPHORIBOSYLAMINOPYRIMIDINE DEAMINASE (EC 3.5.4.26)
DE (RIBOFLAVIN-SPECIFIC DEAMINASE); 5-AMINO-6-(5-
DE PHOSPHORIBOSYLAMINO)URACIL REDUCTASE (EC 1.1.1.193) (HTP REDUCTASE)].
GN RIBD OR RIBG OR RV1409 OR MTCY21B4.26.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae; Mycobacterium.
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
RN [1]
RN SEQUENCE FROM N.A.
RP STRAIN=H37RV;
RX MEDLINE; 98295987.
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squires R., Sulston J.E.,
RA Taylor K., Whitehead S., Barrrell B.G.;
RA "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).
CC -1- FUNCTION: CONVERTS 2,5-DIAMINO-6-(RIBOSYLAMINO)-4(3H)-PYRIMIDINONE
CC 5'-PHOSPHATE INTO 5-AMINO-6-(RIBOSYLAMINO)-2,4(1H,3H)-
CC PYRIMIDINEDIONE 5'-PHOSPHATE.
CC -1- CATALYTIC ACTIVITY: 2,5-DIAMINO-6-HYDROXY-4-(5-
CC PHOSPHORIBOSYLAMINO)PYRIMIDINE + H(2)O = 5-AMINO-6-(5-
CC PHOSPHORIBOSYLAMINO)URACIL + NH(3).
CC -1- CATALYTIC ACTIVITY: 5-AMINO-6-(5-PHOSPHORIBITYLAMINO)URACIL +
CC NADP(+) = 5-AMINO-6-(5-PHOSPHORIBOSYLAMINO)URACIL + NADPH.
CC -1- COFACTOR: ZINC (BY SIMILARITY).
CC -1- PATHWAY: SECOND AN THIRD STEPS OF RIBOFLAVIN BIOSYNTHESIS.
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE CYTIDINE AND
CC DEOXYCYTIDYLATE DEAMINASES FAMILY.
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE HTP
CC REDUCTASE FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; Z80108; CAB02188.1; -
CC TUBERCULIST; RV1409; -
CC INTERPRO; IPR002125; -
CC INTERPRO; IPR002734; -
CC PFAM; PF01872; RIBD_C; 1.
CC PFAM; PF00383; dCMP_cyt_deam; 1.
CC PROSITE; PS00903; CYT_DCMP_DEAMINASES; 1.
CC Ribo flavin biosynthesis; Hydrolase; Zinc; Oxidoreductase; NADP;
KW Multifunctional enzyme.
FT DOMAIN 1 152 DEAMINASE.
FT DOMAIN 153 339 REDUCTASE.
-----
```

```
FT METAL 57 57 ZINC (BY SIMILARITY).
FT METAL 82 82 ZINC (BY SIMILARITY).
FT METAL 91 91 ZINC (BY SIMILARITY).
SQ SEQUENCE 339 AA; 35366 MW; 1DC0A1B8E7EC08B4 CRC64;

Query Match 43.5%; Score 216.5; DB 1; Length 339;
Best Local Similarity 48.9%; Pred. No. 3.6e-17;
Matches 45; Conservative 13; Mismatches 33; Indels 1; Gaps 1;

QY 1 MDVAIQLPNMVQGTGVNPPVGVAVVNEGRIVGIGAHLRKGDHAEVQALDMAQXNAEGA 59
  1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 13 MGLAIEHSYQVKGITTPKPVGAVIVDPNGRIVGAGTEPAGGDHAEVVALRRAGGLAAG 72
  1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 60 ATIVITLPCSHFGSTGTPPCVNKIIDCKIAXV 91
  1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 73 AIVVTHPCNHGKTPPCVNALIEARGVV 104
  1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 6
RIBD_HAEIN
ID RIBD_HAEIN STANDARD; PRT; 372 AA.
AC P44326;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE RIBOFLAVIN BIOSYNTHESIS PROTEIN RIBD [INCLUDES:
DE DIAMINOHYDROXYPHOSPHORIBOSYLAMINOPYRIMIDINE DEAMINASE (EC 3.5.4.26)
DE (RIBOFLAVIN-SPECIFIC DEAMINASE); 5-AMINO-6-(5-
DE PHOSPHORIBOSYLAMINO)URACIL REDUCTASE (EC 1.1.1.193) (HTP REDUCTASE)].
GN RIBD OR RIBG OR H10944.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
RN [1]
RN SEQUENCE FROM N.A.
RP STRAIN=RD / KW20;
RX MEDLINE; 95350630.
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uutterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RA "Whole-genome random sequencing and assembly of Haemophilus
RT influenzae Rd."
RL Science 269:496-512(1995).
CC -1- FUNCTION: CONVERTS 2,5-DIAMINO-6-(RIBOSYLAMINO)-4(3H)-PYRIMIDINONE
CC 5'-PHOSPHATE INTO 5-AMINO-6-(RIBOSYLAMINO)-2,4(1H,3H)-
CC PYRIMIDINEDIONE 5'-PHOSPHATE.
CC -1- CATALYTIC ACTIVITY: 2,5-DIAMINO-6-HYDROXY-4-(5-
CC PHOSPHORIBOSYLAMINO)PYRIMIDINE + H(2)O = 5-AMINO-6-(5-
CC PHOSPHORIBOSYLAMINO)URACIL + NH(3).
CC -1- CATALYTIC ACTIVITY: 5-AMINO-6-(5-PHOSPHORIBITYLAMINO)URACIL +
CC NADP(+) = 5-AMINO-6-(5-PHOSPHORIBOSYLAMINO)URACIL + NADPH.
CC -1- COFACTOR: ZINC (BY SIMILARITY).
CC -1- PATHWAY: SECOND AN THIRD STEPS OF RIBOFLAVIN BIOSYNTHESIS.
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE CYTIDINE AND
CC DEOXYCYTIDYLATE DEAMINASES FAMILY.
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE HTP
CC REDUCTASE FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC
```

	EMBL; U32775; AAC22598.1; -
DR	TIGR: HI0944; -
DR	INTERPRO: IPR002125; -
DR	INTERPRO: IPR002734; -
DR	Pfam: PF01872; RIBD C; 1.
DR	Pfam: PF00383; dCMP_cyt_deam_1;
DR	ProSITE; PS00903; CYT_dCMP_DEAMINASES; 1.
KW	Riboflavin biosynthesis; Hydrolase; Zinc; Oxidoreductase; NADP;
MU	Multifunctional enzyme.
FT	DMAIN 1 150 DEAMINASE.
FT	DMAIN 151 372 REDUCTASE.
FT	METAL 55 55 ZINC (BY SIMILARITY).
FT	METAL 80 80 ZINC (BY SIMILARITY).
FT	METAL 89 89 ZINC (BY SIMILARITY).
SQ	SEQUENCE 372 AA; 40931 MW; C7504EDDBB86F7D CRC64;

  

Query Match            43.0%; Score 214; DB 1; Length 372;  
Best Local Similarity 47.3%; Pred. No. 7.6e-17;  
Matches 43; Conservative 11; Mismatches 37; Indels 0; Gaps 0;

  

QY	1 MDVAIOLPNMVCGXTGVNPPVGAVVNEGRIVGIGHLRKGDHAEVQAOLDMAOXNAEGA 60
Db	12 MQRALDLAAKGQVTTPNPNSVCGLVKNGEIVEGGFHFKAGQPFAERVALAQAGENAKGA 71
	:   :   :   :   :   :   :   :   :   :   :   :   :   :   :
QY	61 TIYTLEPCSHFGSTPTCPVNKIIDCKIAKV 91
Db	72 TAVTVLPCAHYGRTPPALGLEAGVKKVI 102
	: ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  :

  

```
RESULT 7
RIBD_SYNY3 STANDARD; PRT; 368 AA.
AC Q55158;
ID RIBD_SYNX3
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE RIBOFLAVIN BIOSYNTHESIS PROTEIN RIBD [INCLUDES:
DE DIAMINO-HYDROXYPHOSPORIBOSYLAMINOPYRIDINE DEAMINASE (EC 3.5.4.26)
DE (RIBOFLAVIN-SPECIFIC DEAMINASE); 5'-AMINO-6-(5-
DE PHOSPORIBOSYLAMINO)URACIL REDUCTASE (EC 1.1.1.193) (HTP REDUCTASE)].
GN RIBD OR SLR0066.
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
[1]
RN RP SEQUENCE FROM N.A.
RX MEDLINE; 96127529.
RA Kaneko T., Tanaka A., Sato S., Kotani H., Suzuki T., Miyajima N.,
RA Sugiyura M., Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
RT region from map positions 64% to 92% of the genome.";
RL DNA Res. 2:153-166(1995).
CC -|- FUNCTION: CONVERTS 2,5-DIAMINO-6-(RIBOSYLAMINO)-4(3H)-PYRIMIDINONE
CC 5'-PHOSPATE INTO 5'-AMINO-6-(RIBOSYLAMINO)-2,4(LH,3H)-
CC PYRIMIDINEDIONE 5'-PHOSPATE.
CC -|- CATALYTIC ACTIVITY: 2,5-DIAMINO-6-HYDROXY-4-(5-
CC PHOSPORIBOSYLAMINO)PYRIMIDINE + H(2)O = 5-AMINO-6-(5-
CC PHOSPORIBOSYLAMINO)URACIL + NH(3).
CC -|- CATALYTIC ACTIVITY: 5-AMINO-6-(5-PHOSFORIBOSYLAMINO)URACIL +
CC NADP(+)= 5-AMINO-6-(5-PHOSPORIBOSYLAMINO))URACIL + NADPH.
CC -|- COFACTOR: ZINC (BY SIMILARITY).
CC -|- PATHWAY: SECOND AN THIRD STEPS OF RIBOFLAVIN BIOSYNTHESIS.
CC -|- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE CYTIDINE AND
CC DEOXYCYTIDYLATE DEAMINASES FAMILY.
CC -|- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE HTP
CC REDUCTASE FAMILY.
-----
This SWISS-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
institutions without permission of the Centre for Molecular Biology and Biotechnology Ltd. is prohibited.
```



CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR	EMBL; D26185; BAA05254.1; -			
DR	EMBL; Z99104; CAB11794.1; -			
DR	PIR; S11690; S11690.			
DR	SUBTILIST; BG10082; YAAJ.			
DR	INTERPRO; IPR002125; -			
DR	PFAM; PF00383; dCMP_cyt_deam; 1.			
DR	PROSITE; PS00903; Cyt_DCMP_DEAMINASES; 1.			
KW	Hypothetical protein; Hydrolase; Zinc.			
FT	METAL 53 53	ZINC (BY SIMILARITY).		
FT	METAL 83 83	ZINC (BY SIMILARITY).		
FT	METAL 86 86	ZINC (BY SIMILARITY).		
SQ	SEQUENCE 161 AA; 17751 MW; 3256F31DF6610FBB CRC64;			

```
Query Match      23.6%; Score 117.5; DB 1; Length 161;  
Best Local Similarity 33.3%; Pred. No. 2.3e-06;  
Matches         27; Conservative 16; Mismatches 23; Indels 15; Gaps 3;  
  
QY    20 PVGAVVVNEGRIVCIGAHRLRGDK-----HAEVQLD----MAOXNNEGATIIYTLPSCS 70  
       |::|::|| :: ||:: :|||: ||   ; |||:::|||  
Db     25 PIGAVLVINGEITARAHLNRETEQRSTAAHEMLVIDEACRALGTWRLEGATLYVTLEPC- 83  
  
QY    71 HFGSTPPCVNKIIDCKTAXVV 91  
       | | :: :| ||  
Db     84 -----PMCAGAVVLRSRVEKVV 99
```

RESULT	12
YFHC_HAEIN	
ID	YFHC_HAEIN      STANDARD;      PRT;    173 AA.
AC	P44931;
DT	01-NOV-1995 (Rel. 32, Created)
DT	01-NOV-1995 (Rel. 32, Last sequence update)
DT	01-NOV-1995 (Rel. 32, Last annotation update)
DE	HYPOTHETICAL PROTEIN HI0906.
GN	HI0906.
OS	Haemophilus influenzae.
OC	Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC	Haemophilus.
RN	[1]
RC	SEQUENCE FROM N.A.
RP	STRAIN=RD / KW20;
RX	MEDLINE: 95350630.
RA	Fleischmann R.B., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA	Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA	McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA	Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA	Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA	Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA	Fine L.D., Frichman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA	Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA	Venter J.C.;
RT	"whole-genome random sequencing and assembly of Haemophilus
RT	Influenzae Rd.";
RL	Science 269:496-512(1995).
CC	- - COFACTOR: ZINC (BY SIMILARITY).
CC	- - SIMILARITY: BELONGS TO THE CYTIDINE AND DEOXYCYTIDYLATE DEAMINASES
CC	FAMILY. STRONG, TO E.COLI AND B.SUBTILIS RIBG. STRONG, TO E.COLI
CC	YFHC.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/announcement/or\\_send\\_an\\_email\\_to\\_license@isb-sib.ch](http://www.isb-sib.ch/announcement/or_send_an_email_to_license@isb-sib.ch)).

```
DR EMBL; U32772; AAC22565.1; -.
DR TIGR; H10906; -.
DR INTERPRO; IPR002125; -.
DR PFAM; PF00383; dCMP_cyt_deam; 1.
DR PROSITE; PS00903; CYT_DCMP_DEAMINASES; 1.
KW Hypothetical protein; Hydrolase; Zinc.
FT METAL 61 61 ZINC (BY SIMILARITY).
FT METAL 91 91 ZINC (BY SIMILARITY).
FT METAL 94 94 ZINC (BY SIMILARITY).
SQ SEQUENCE 173 AA; 19359 MW; 785B2CE532FB72DE CRC64;

Query Match 21.5%; Score 107; DB 1; Length 173;
Best Local Similarity 32.7%; Pred. No. 3.8e-05;
Matches 33; Conservative 20; Mismatches 30; Indels 18; Gaps 6;

QY 1 MDYAIQLPNVVOGTGVNPPGAVVNEGR-IVGIGAHLR--KCD--KHAEVQALDMAQX 55
| | | | | : : | | | | | : : | | | : : | | | | |
15 MRYALELADRAEALGETI--PVGAVLYDDARNIIGEGWNLIVSDSDPTAHAEIILNRGAK 72

QY 56 NAE-----GATVITLPCSHFGSTPPCVNKLIDCKIAXVV 91
| : : : | | | | | : | : : | | : | : |
73 NIQYRLNLTLYLLEPCT-----MCAGAILHSRIKRLV 107

Db 73 NIQYRLNLTLYLLEPCT-----MCAGAILHSRIKRLV 107

RESULT 13
YFHC_ECOLI
ID YFHC_ECOLI STANDARD; PRT; 178 AA.
AC P30134;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE HYPOTHETICAL 20.0 KDA PROTEIN IN PURL-DPJ INTERGENIC REGION (ORF178).
GN YFHC.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NWL37;
RX MEDLINE; 92292954.
RA Poulsen L.K., Larsen N.W., Molin S., Andersson P.;
RT "Analysis of an Escherichia coli mutant strain resistant to the cell-
killing function encoded by the gef gene family.";
RL Mol. Microbiol. 6:895-905(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RA Nashimoto H., Salto N.;
RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE; 97426617.
RA Blatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1433-1474(1997).
CC -!- FUNCTION: MUTATION IN THIS PROTEIN MAKES E. COLI RESISTANT TO THE
TOXIC PROTEINS ENCODED BY THE GEF GENE FAMILY.
CC -!- COFACTOR: ZINC (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE CYTIDINE AND DEOXYCYTIDYLATE DEAMINASES
FAMILY. STRONG. TO E. COLI AND B. SUBTILIS RIBG. STRONG, TO
H. INFLUENZAE H10906.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
CC EMBL; A5000713; AAC07025.1; -.
DR INTERPRO; IPR002125; -.
DR PFAM; PF00383; dCMP_cyt_deam; 1.
DR PROSITE; PS00903; CYT_DCMP_DEAMINASES; 1.
KW Hypothetical protein; Hydrolase; Zinc.
FT METAL 52 52 ZINC (BY SIMILARITY).
FT METAL 82 82 ZINC (BY SIMILARITY).
```

```
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; X72336; CAA51064.1; -.
DR EMBL; D64044; BAA10909.1; -.
DR EMBL; U36841; AAA79821.1; -.
DR EMBL; AE000342; AAC75612.1; -.
DR PIR; S20974; S20974.
DR ECOGENE; EG11372; YFHC.
DR INTERPRO; IPR002125; -.
DR PFAM; PF00383; dCMP_cyt_deam; 1.
DR PROSITE; PS00903; CYT_DCMP_DEAMINASES; 1.
KW Hypothetical protein; Hydrolase; Zinc.
FT METAL 68 68 ZINC (BY SIMILARITY).
FT METAL 98 98 ZINC (BY SIMILARITY).
FT METAL 101 101 ZINC (BY SIMILARITY).
FT MUTAGEN 64 64 D->E: RESISTANCE TO THE CELL-KILLING
FUNCTION ENCODED BY THE GEF GENE FAMILY.
SQ SEQUENCE 178 AA; 20026 MW; 80B2E3B5FD61AA8A CRC64;

Query Match 21.2%; Score 105.5; DB 1; Length 178;
Best Local Similarity 36.1%; Pred. No. 5.7e-05;
Matches 30; Conservative 9; Mismatches 25; Indels 19; Gaps 4;

QY 20 PVGAVVNEGRIVG-----IGAHLRKDGKHAEVQALD-----MAQXNAEGATYITLEP 68
| | | | | : : | | | | | : : | | | | |
40 PVGAVLVHNNRVIGEGWNRPIGRH--DPTAHAEIEMALRGGLVMQNYRLIDATLYVTLEP 97

QY 69 CSHFGSTPPCVNKLIDCKIAXVV 91
| : : : | | | | | : | : : | | : | : |
98 C-----VMCAGAILHSRIGRVV 114

Db 98 C-----VMCAGAILHSRIGRVV 114

RESULT 14
Y903_AQUAE
ID Y903_AQUAE STANDARD; PRT; 151 AA.
AC O67050;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE HYPOTHETICAL PROTEIN AQ_903.
GN AQ_903.
OS Aquifex aeolicus.
OC Bacteria; Aquificales; Aquificaceae; Aquifex.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VF5;
RX MEDLINE; 98196666.
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Snead M.A., Keller M., Aujay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
aeolicus.";
RL Nature 392:353-358(1998).
CC -!- COFACTOR: ZINC (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE CYTIDINE AND DEOXYCYTIDYLATE DEAMINASES
FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
CC EMBL; A5000713; AAC07025.1; -.
DR INTERPRO; IPR002125; -.
DR PFAM; PF00383; dCMP_cyt_deam; 1.
DR PROSITE; PS00903; CYT_DCMP_DEAMINASES; 1.
KW Hypothetical protein; Hydrolase; Zinc.
FT METAL 52 52 ZINC (BY SIMILARITY).
FT METAL 82 82 ZINC (BY SIMILARITY).
```

Job time: 177 sec

```
FT METAL      85      85      ZINC (BY SIMILARITY).
SQ SEQUENCE 151 AA; 17246 MW; 5543DC00B9846D25 CRC64;

Query Match      19.3%; Score 96; DB 1; Length 151;
Best Local Similarity 35.4%; Pred. No. 0.00056;
Matches 29; Conservative 10; Mismatches 19; Indels 24; Gaps 5;

QY 20 PVGAVVNEGRIVGIGAH-----LRKGDKHAEVQALDMA--QXNA---EGATYYITLPEPC 69
   ||||:| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 24 PVGAIIIVKEGELIS-KAHNSVEELKDPTAHAEMLAIKEACRLNTKYLEGCGLYVTLPEPC 82

QY 70 SHEGSTPPCVNKIIIDCKIAXV 91
   | | | | |
Db 83 -----IMCSYALVL 91
```

```
RESULT 15
YILXG_VIBFI
ID YLXG_VIBFI STANDARD; PRT; 147 AA.
AC P33968;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last annotation update)
DE HYPOTHETICAL 16.6 KDA PROTEIN IN LUXG 3 REGION.
OS Vibrio fischeri.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 7744;
RX MEDLINE; 93320098.
RA Lee C.Y., Zittner R.B., Miyamoto C.M., Meighen E.A.;
RT "The gene convergent to luxG in Vibrio fischeri codes for a protein
RL Biochim. Biophys. Acta 1143:337-339(1993).
CC -!- COFACTOR: ZINC (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE CYTIDINE AND DEOXYCYTIDYLATE DEAMINASES
CC FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; X70289; CAA49769.1; -.
DR INTERPRO; IPR002125; -.
DR PFAM; PF00383; dCMP_cyt_deam; 1.
DR PROSITE; PS00903; CYT_DCMP_DEAMINASES; 1.
KW Hypothetical protein; Hydrolase; Zinc.
FT METAL 67 67 ZINC (BY SIMILARITY).
FT METAL 92 92 ZINC (BY SIMILARITY).
FT METAL 95 95 ZINC (BY SIMILARITY).
SQ SEQUENCE 147 AA; 16622 MW; 174FDA8DC1458D49 CRC64;
```

```
Query Match      17.8%; Score 88.5; DB 1; Length 147;
Best Local Similarity 30.3%; Pred. No. 0.0038;
Matches 27; Conservative 10; Mismatches 25; Indels 27; Gaps 3;

QY 21 VGAVVNEGRIVGIG-----AHLRKGDKHAEVQALDMAQXNAEGAT 61
   ||||:| | | | | | | | | | | | | | | | | | | | | | | | |
Db 27 VGAVITKINRIVSGFNGPHGVSDSADTDREIKYLK--TLHAEENAILFAKRDLEGCD 84

QY 62 IYITLPCSHFGSTPPCVNKIIIDCKIAXV 90
   | : | | | | | | | | | | | | | | | | | | | | | | | | |
Db 85 IWTHTFPC-----PNCAAKIIQTGISKV 107
```

**THIS PAGE BLANK (USPTO)**



Result No.	Query			Length	DB	ID	Description
	Score	Match					
1	229	46.0	245	2	O51825		O51825 shigella fl
2	213	42.8	357	2	Q9RY01		Q9RY01 deinococcus
3	213	42.8	363	10	Q9SUB7		Q9SUB7 arabidopsis
4	211.5	42.5	363	2	Q9RE66		Q9RE66 bartonella
5	210.5	42.3	348	2	Q9X2B8		Q9X2B8 thermotoga
6	194	39.0	241	2	O86961		O86961 thermotoga
7	191	38.4	375	2	O84735		O84735 chlamydia t
8	113.5	22.8	405	3	P87241		P87241 schizosacch
9	110.5	22.2	145	2	Q9X3V3		Q9X3V3 pseudomonas
10	107	21.5	164	2	P73717		P73717 synechocyst
11	105.5	21.2	376	2	Q9RKM1		Q9RKM1 streptomyce
12	105	21.1	600	2	Q9RV23		Q9RV23 deinococcus
13	104.5	21.0	1307	10	Q9S710		Q9S710 arabidopsis
14	101	20.3	145	2	Q9RR66		Q9RR66 deinococcus
15	99	19.9	599	10	Q9STV4		Q9STV4 arabidopsis
16	92.5	18.6	201	2	Q9WYN5		Q9WYN5 thermotoga
17	89.5	18.0	150	2	O51833		O51833 porphyromon
18	88.5	17.8	124	2	Q9S3V8		Q9S3V8 vibrio fisc
19	85.5	17.2	152	2	O69719		O69719 mycobacteri

```
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE RIBOFLAVIN-SPECIFIC DEAMINASE.
GN DR0153.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RL;
RX MEDLINE; 20036896.
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome Sequence of the Radioreistant Bacterium Deinococcus
RT radiodurans RL.";
RL Science 286:1571-1577(1999).
RL EMBL: AE001878; AAF09742.1; -.
DR TIGR: DR0153; -.
DR INTERPRO: IPR002125; -.
DR INTERPRO: IPR002734; -.
DR PFAM: PF00383; dCMP_cyt_deam; 1.
DR PFAM: PF01872; RibD_C; 1.
DR PROSITE: PS00903; CYT_DCMP_DEAMINASES; 1.
SQ SEQUENCE 357 AA; 37226 MW; ADBFE65C10626007 CRC64;

Query Match 42.8%; Score 213; DB 2; Length 357;
Best Local Similarity 52.4%; Pred. No. 2.8e-16;
Matches 44; Conservative 9; Mismatches 27; Indels 4; Gaps 1;

QY 13 GXTGVNPPVGVVNEG-----RVIGTGAHLRKGDKHAQVQALDMAQXNAEGATYITLPEP 68
Db 24 GRTSPNPPVGVVIRDEGIIASEIVGRGPHKAGEPFAEVALRAGERARGATYITLPEP 83
QY 69 CSHFGSTPPCVNKIKDCKIAKXVVL 92
Db 84 CSHDGRTPPCADALIAGVAVVV 107

RESULT 3
Q9SUB7 PRELIMINARY; PRT; 363 AA.
AC Q9SUB7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE HYPOTHETICAL 39.7 KDA PROTEIN.
GN T13K14.120.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; Rosidae; eurosids II; Brassicales;
OC Brassicaceae; Arabidopsis.
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Pohl T., Weizenegger T., Bancroft I., Mewes H.W.,
RA Mayer K.F.X., Lemcke K., Schueller C.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL080282; CAB45891.1; -.
DR INTERPRO: IPR002125; -.
DR PFAM: PF00383; dCMP_cyt_deam; 1.
DR PROSITE: PS00903; CYT_DCMP_DEAMINASES; UNKNOWN_1.
KW Hypothetical protein
SQ SEQUENCE 363 AA; 39729 MW; 1281410D8FE272F6 CRC64;
```

```
Query Match 42.8%; Score 213; DB 10; Length 363;
Best Local Similarity 46.7%; Pred. No. 2.9e-16;
Matches 43; Conservative 12; Mismatches 37; Indels 0; Gaps 0;

QY 1 MDVAIQLPNMVQGXGTGVNPPVGVVNEGRIVGIGAHLRKGDKHAQVQALDMAQXNAEGA 60
Db 15 MRKCVELAKRAIGTSPNPMVGCIVKDGDIVGQGFHPKAGQPHAEVFAIRDAGELAENA 74
QY 61 TIVITLPCSHFGSTPPCVNKIKDCKIAKXVVL 92
Db 75 TAYVLEPCNHYGRTPTCTEALINAKVRRVVI 106

RESULT 4
Q9REF6 PRELIMINARY; PRT; 363 AA.
ID Q9REF6;
AC Q9REF6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE RIBOFLAVIN-SPECIFIC DEAMINASE / REDUCTASE (FRAGMENT).
GN RIBD.
OS Bartonella henselae (Rochalimaea henselae).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Bartonellaceae; Bartonella.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HOUSTON-1;
RA Bereswill S., Hinkelmann S., Kist M., Sander A.;
RT "Molecular analysis of riboflavin synthesis genes in bartonella
RT henselae and use of the ribC gene for differentiation of bartonella
RT species by PCR.";
RL J. Clin. Microbiol. 37:3159-3166(1999).
DR EMBL: AJ132928; CAB63090.1; -.
DR INTERPRO: IPR000886; -.
DR INTERPRO: IPR002125; -.
DR INTERPRO: IPR002734; -.
DR PFAM: PF00383; dCMP_cyt_deam; 1.
DR PFAM: PF01872; RibD_C; 1.
DR PROSITE: PS00014; ER_TARGET; UNKNOWN_1.
DR PROSITE: PS00903; CYT_DCMP_DEAMINASES; UNKNOWN_1.
FT NON_TER 1
SQ SEQUENCE 363 AA; 39524 MW; 131630AF56074F7D CRC64;

Query Match 42.5%; Score 211.5; DB 2; Length 363;
Best Local Similarity 47.9%; Pred. No. 4.3e-16;
Matches 45; Conservative 10; Mismatches 34; Indels 5; Gaps 1;

QY 4 AIQLPNMVQGXGTGVNPPVGVVNEG-----IVGIGAHLRKGDKHAQVQALDMAQXNAE 58
Db 2 AIRLAERHVLGTGENSVGTIIARNDENVCVYIVGVGTATQGRPHAEVQALOMAGSLAH 61
QY 59 GATYITLPCSHFGSTPPCVNKIKDCKIAKXVVL 92
Db 62 GATAYVTLPSCSHYGTSPCVNTLLKSGISRVI 95

RESULT 5
Q9X2E8 PRELIMINARY; PRT; 348 AA.
ID Q9X2E8;
AC Q9X2E8;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE RIBOFLAVIN-SPECIFIC DEAMINASE.
GN TM1828.
OS Thermotoga maritima.
OC Bacteria; Thermotogales; Thermotoga.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MSB8 / DSM 3109;
```



[illegible]

RL	Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.					
DR	EMBL:	AC011914;	AF15073.1;	-		
DR	EMBL:	AC008075;	AAD49971.1;	-		
DR	INTERPRO:	IPR002125;	-			
DR	PFAM:	PF00383;	dCMP_cyt_deam;	1.		
SQ	SEQUENCE 1307 AA; 8CDAD95ABAB24719 CRC64;					
		Query Match	21.0%;	Score 104.5;	DB 10;	Length 1307;
		Best Local Similarity	31.0%;	Pred. No. 0.0025;		
Matches	27;	Conservative	14;	Mismatches	19;	Indels 27; Gaps
OY	20	PVGAVVNVNIGRIVGIG-----AHLRGDKHAEVOALDMAOXNAEGA-----TIIYI 64      : : :           :   :				
Db	1131	PVGAVLVHDGKIIFARGYNLVEELRSTAHAEICI-----REGSKALRSWRLADTTLYV 1184      : : :           :   :				
OY	65	TLEPCSHFGSTPPCVNKNIIDCKIAVV 91               : :: :				
Db	1185	TLEPC-----PMCAGAILQARVNTLV 1205               : :: :				
RESULT	14					
Q9RR66	ID	Q9RR66	PRELIMINARY;	PRT;	145 AA.	
AC	Q9RR66;					
DT	01-MAY-2000	(TREMBLrel. 13,	Created)			
DT	01-MAY-2000	(TREMBLrel. 13,	Last sequence update)			
DE	01-JUN-2000	(TREMBLrel. 14,	Last annotation update)			
DE	DEOXYCYTIDYLATE DEAMINASE,	PUTATIVE.				
GN	DR2631.					
OS	Deinococcus radiodurans.					
OC	Bacteria; Thermus/Deinococcus group;	Deinococcales;	Deinococcus.			
CC	[1]					
RN	SEQUENCE FROM N.A.					
RP	STRAIN=RL;					
RC	WHITE O.; Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,					
RA	White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,					
RA	Dodson R.J., Qian D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,					
RA	Moffat K.S., Haft H., Jiang L., Pamphile W., Crosby M., Shen M.,					
RA	Vamathevan J.J., Lam P., McDonald L., Utterback T., Zaleski C.,					
RA	Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,					
RA	Ketchum K.A., Nelson K.E., Saizberg S., Smith H.O., Venter J.C.,					
RA	Fraser C.M.;					
RT	"Genome Sequence of the Radioresistant Bacterium Deinococcus					
RT	radiodurans RL."					
RL	Science 286:1571-1577(1999).					
DR	EMBL:	AE002092;	AAF12167.1;	-		
DR	TIGR:	DR2631;	-			
DR	INTERPRO:	IPR002125;	-			
DR	PFAM:	PF00383;	dCMP_cyt_deam;	1.		
DR	PROSITE:	PS00903;	CYT_DCMP_DEAMINASES;	UNKNOWN_1.		
SQ	SEQUENCE 145 AA; 15547 MW; 110CA05197E999A CRC64;					
		Query Match	20.3%;	Score 101;	DB 2;	Length 145;
		Best Local Similarity	34.1%;	Pred. No. 0.00046;		
Matches	28;	Conservative	11;	Mismatches	21;	Indels 22; Gaps
OY	21	VGAVVNV-NGRIVGIGAHLR-----KDGKHAEVQALDMAOXNAEGATIIYI 64      : : :           :   :				
Db	28	VGACILDRHRHVGVGYNGRAAGEPNRESLAQGASGYTHAEVNALLAANNNGEGHTLYV 87      :~ :~ :~          :   :				
OY	65	TLEPCSHFGSTPPCVNKNIIDCK 86               : :: :				
Db	88	THEPCS-----VCARLIVNSR 103               : :: :				
RESULT	15					
Q9STY4	ID	Q9STY4	PRELIMINARY;	PRT;	599 AA.	
AC	Q9STY4;					
DT	01-MAY-2000	(TREMBLrel. 13,	Created)			
DT	01-MAY-2000	(TREMBLrel. 13,	Last sequence update)			

Search completed: November 14, 2000, 10:10:10  
Job time: 207 sec